

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2004, 13:44:42 ; Search time 16.5 Seconds

(without alignments)
1339.146 Million cell updates/sec

Title: US-10-091-085-3

Perfect score: 2250

Sequence: 1 MATSWGTVFPMVVCVCSA.....ETGWAIGATPHLLQSLGISH 428

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgm2_6/ptodata/2/iaa/5A.COMB.pep.*
- 2: /cgm2_6/ptodata/2/iaa/5B.COMB.pep.*
- 3: /cgm2_6/ptodata/2/iaa/6A.COMB.pep.*
- 4: /cgm2_6/ptodata/2/iaa/6B.COMB.pep.*
- 5: /cgm2_6/ptodata/2/iaa/PCURS.COMB.pep.*
- 6: /cgm2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2250	100.0	428	4	US-09-608-285A-3
2	2250	100.0	428	4	US-09-608-285A-5
3	2250	100.0	428	4	US-09-240-639-6
4	2250	100.0	428	4	US-09-240-639-9
5	2250	100.0	428	4	US-09-350-836B-3
6	2250	100.0	428	4	US-09-350-836B-5
7	2250	100.0	428	4	US-09-370-265-3
8	2250	100.0	428	4	US-09-370-265-5
9	2250	100.0	428	4	US-09-557-800C-3
10	2250	100.0	428	4	US-09-557-800C-5
11	2250	100.0	428	4	US-09-370-625A-3
12	2250	100.0	428	4	US-09-370-625A-5
13	2235	99.3	428	4	US-09-608-285A-7
14	2235	99.3	428	4	US-09-350-836B-7
15	2235	99.3	428	4	US-09-370-265-7
16	2235	99.3	428	4	US-09-557-800C-7
17	2235	99.3	428	4	US-09-370-625A-7
18	2104	93.5	405	4	US-09-608-285A-25
19	2104	93.5	405	4	US-09-370-265-25
20	2104	93.5	405	4	US-09-557-800C-25
21	2104	93.5	405	4	US-09-370-625A-25
22	1837.5	81.7	465	4	US-09-557-800C-56
23	1837.5	81.7	465	4	US-09-370-625A-39
24	1832.5	81.4	465	4	US-09-240-639-8
25	999	44.4	456	4	US-09-240-639-2
26	999	44.4	484	4	US-09-608-285A-27
27	999	44.4	484	4	US-09-370-265-27

28	999	44.4	484	4	US-09-557-800C-27	Sequence 27, Appl
29	999	44.4	484	4	US-09-370-625A-27	Sequence 27, Appl
30	819.5	36.4	471	4	US-09-608-285A-60	Sequence 60, Appl
31	507.5	22.6	459	4	US-09-129-112-9	Sequence 9, Appl
32	503.5	22.4	467	4	US-09-129-112-19	Sequence 19, Appl
33	498	22.1	462	4	US-09-129-112-2	Sequence 2, Appl
34	491.5	21.8	462	4	US-09-129-112-15	Sequence 15, Appl
35	489	21.7	455	4	US-09-240-639-10	Sequence 10, Appl
36	467.5	20.8	473	4	US-09-240-639-12	Sequence 12, Appl
37	446.5	19.8	454	4	US-09-240-639-11	Sequence 11, Appl
38	351	15.6	502	4	US-09-557-800C-55	Sequence 55, Appl
39	351	15.6	502	4	US-09-370-625A-38	Sequence 38, Appl
40	351	15.6	510	3	US-08-930-921-1	Sequence 1, Appl
41	347.5	15.4	529	4	US-09-240-639-4	Sequence 4, Appl
42	234	10.4	148	4	US-09-240-639-17	Sequence 17, Appl
43	186	8.3	153	4	US-09-240-639-13	Sequence 13, Appl
44	180.5	8.0	150	4	US-09-240-639-16	Sequence 16, Appl
45	179	8.0	153	4	US-09-240-639-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1

US-09-608-285A-3
; Sequence 3, Application US/09608285A

; Patent No. 6335013

; GENERAL INFORMATION:

; APPLICANT: Ford, John

; APPLICANT: Mulero, Julio

; APPLICANT: Yeung, George

; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE

; TITLE OF INVENTION: POLYPEPTIDES

; FILE REFERENCE: 28110/36570

; CURRENT APPLICATION NUMBER: US/09/608,285A

; CURRENT FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: 09/583,231

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 09/557,800

; PRIOR FILING DATE: 2000-04-25

; PRIOR APPLICATION NUMBER: 09/481,238

; PRIOR FILING DATE: 2000-01-11

; PRIOR APPLICATION NUMBER: 09/370,265

; PRIOR FILING DATE: 1999-08-09

; PRIOR APPLICATION NUMBER: PCT/US99/16180

; PRIOR FILING DATE: 1999-07-16

; PRIOR APPLICATION NUMBER: 09/350,836

; PRIOR FILING DATE: 1999-07-09

; PRIOR APPLICATION NUMBER: 09/273,447

; PRIOR FILING DATE: 1999-03-19

; PRIOR APPLICATION NUMBER: 09/244,444

; PRIOR FILING DATE: 1999-02-04

; PRIOR APPLICATION NUMBER: 09/122,449

; PRIOR FILING DATE: 1998-07-24

; PRIOR APPLICATION NUMBER: 09/118,205

; PRIOR FILING DATE: 1998-07-16

; NUMBER OF SEQ ID NOS: 60

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3

; LENGTH: 428

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-608-285A-3

Query Match

Best Local Similarity 100.0%; Score 2250; DB 4; Length 428;

Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATSWGTVFPMVVCVCSAVSHRNQQTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Db 1 MATSWGTVFPMVVCVCSAVSHRNQQTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy 61 RIHYTFVQKMPQLPIEGEVFDSVVKPGLSAFVDQPKQAETVQGLLEVAKDSIPRSHW 120

Db

61

RIHVYTFVQKMPGOLPILEGEVDSVKPGLSAFVDPQKGAETVQGLLEVAKDSIPRSHW

120

Qy

121

KKTPVVLKATAGLRLLPEHKAALLFEVKEIFPKSPFLVPKGSVINDGDEGLAWVTV

180

Db

121

KKTPVVLKATAGLRLLPEHKAALLFEVKEIFPKSPFLVPKGSVINDGDEGLAWVTV

180

Qy

181

NFLTQGLHGHROETVGTLDLGASTQITFLPQFEXTLEQTPRGYLTSEFMFNSTYKLYTH

240

Db

181

NFLTQGLHGHROETVGTLDLGASTQITFLPQFEXTLEQTPRGYLTSEFMFNSTYKLYTH

240

Qy

241

SYLGFGLKAARLATLGALETEGDTGHTFRSACLPRWLEAEWIFPGVKYQYGGNQBGEVGF

300

Db

241

SYLGFGLKAARLATLGALETEGDTGHTFRSACLPRWLEAEWIFPGVKYQYGGNQBGEVGF

300

Qy

301

EPCYAEVLRVVRGKLHQPVEVQSGFYAFSYYYDRAVDTMDIDYKGGILKVEDPERKAR

360

Db

301

EPCYAEVLRVVRGKLHQPVEVQSGFYAFSYYYDRAVDTMDIDYKGGILKVEDPERKAR

360

Qy

361

EVCNLENFTSGSPFLCWLDSYITALLKDGFGFADSTVLQTKKNNIETGWLGAATFHL

420

Db

361

EVCNLENFTSGSPFLCWLDSYITALLKDGFGFADSTVLQTKKNNIETGWLGAATFHL

420

Qy

421

LQSLGISH 428

Db

421

LQSLGISH 428

RESULT 2

US-09-608-285A-5

Sequence 5, Application US/09608285A

Patent No. 6335013

GENERAL INFORMATION:

APPLICANT: Ford, John

APPLICANT: Molero, Julio

TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE

TITLE OF INVENTION: POLYPEPTIDES

FILE REFERENCE: 28110/36570

CURRENT APPLICATION NUMBER: US/09/608,285A

CURRENT FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: 09/583,231

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 09/557,800

PRIOR FILING DATE: 2000-04-25

PRIOR APPLICATION NUMBER: 09/481,238

PRIOR FILING DATE: 2000-01-11

PRIOR APPLICATION NUMBER: 09/370,265

PRIOR FILING DATE: 1999-08-09

PRIOR APPLICATION NUMBER: PCT/US99/16180

PRIOR FILING DATE: 1999-07-16

PRIOR APPLICATION NUMBER: 09/350,836

PRIOR FILING DATE: 1999-07-09

PRIOR APPLICATION NUMBER: 09/273,447

PRIOR FILING DATE: 1999-03-19

PRIOR APPLICATION NUMBER: 09/244,444

PRIOR FILING DATE: 1999-02-04

PRIOR APPLICATION NUMBER: 09/122,449

PRIOR FILING DATE: 1998-07-24

PRIOR APPLICATION NUMBER: 09/118,205

PRIOR FILING DATE: 1998-07-16

NUMBER OF SEQ ID NOS: 60

SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 5

LENGTH: 428

TYPE: PRT

ORGANISM: Homo sapiens

US-09-608-285A-5

Query Match

100.0%; Score 2250; DB 4; Length 428;

Best Local Similarity

100.0%; Pred. No. 6.4e-248;

Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

1

MATSWGTVPFVFLVVCVCSAVSHRNOQTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT

60

Db

1

MATSWGTVPFVFLVVCVCSAVSHRNOQTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT

60

Qy

61

RIHVYTFVQKMPGOLPILEGEVDSVKPGLSAFVDPQKGAETVQGLLEVAKDSIPRSHW

120

Db

61

RIHVYTFVQKMPGOLPILEGEVDSVKPGLSAFVDPQKGAETVQGLLEVAKDSIPRSHW

120

Qy

121

KKTPVVLKATAGLRLLPEHKAALLFEVKEIFPKSPFLVPKGSVINDGDEGLAWVTV

180

Db

121

KKTPVVLKATAGLRLLPEHKAALLFEVKEIFPKSPFLVPKGSVINDGDEGLAWVTV

180

Qy

181

NFLTQGLHGHROETVGTLDLGASTQITFLPQFEXTLEQTPRGYLTSEFMFNSTYKLYTH

240

Db

181

NFLTQGLHGHROETVGTLDLGASTQITFLPQFEXTLEQTPRGYLTSEFMFNSTYKLYTH

240

Qy

241

SYLGFGLKAARLATLGALETEGDTGHTFRSACLPRWLEAEWIFPGVKYQYGGNQBGEVGF

300

Db

241

SYLGFGLKAARLATLGALETEGDTGHTFRSACLPRWLEAEWIFPGVKYQYGGNQBGEVGF

300

Qy

301

EPCYAEVLRVVRGKLHQPVEVQSGFYAFSYYYDRAVDTMDIDYKGGILKVEDPERKAR

360

Db

301

EPCYAEVLRVVRGKLHQPVEVQSGFYAFSYYYDRAVDTMDIDYKGGILKVEDPERKAR

360

Qy

361

EVCNLENFTSGSPFLCWLDSYITALLKDGFGFADSTVLQTKKNNIETGWLGAATFHL

420

Db

361

EVCNLENFTSGSPFLCWLDSYITALLKDGFGFADSTVLQTKKNNIETGWLGAATFHL

420

Qy

421

LQSLGISH 428

Db

421

LQSLGISH 428

RESULT 3

US-09-240-639-6

Sequence 5, Application US/09240639

Patent No. 6350447

GENERAL INFORMATION:

APPLICANT: Chadwick, Brian Paul

APPLICANT: Frischauf, Anna-Maria

TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO CD39-LIKE

TITLE OF INVENTION: POLYPEPTIDES AND NUCLEIC ACIDS

FILE REFERENCE: 9598-066

CURRENT APPLICATION NUMBER: US/09/240,639

CURRENT FILING DATE: 1998-01-29

NUMBER OF SEQ ID NOS: 29

SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 6

LENGTH: 428

TYPE: PRT

ORGANISM: Homo sapiens

US-09-240-639-6

Query Match

100.0%; Score 2250; DB 4; Length 428;

Best Local Similarity

100.0%; Pred. No. 6.4e-248;

Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

1

MATSWGTVPFVFLVVCVCSAVSHRNOQTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT

60

Db

1

MATSWGTVPFVFLVVCVCSAVSHRNOQTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT

60

Qy

61

RIHVYTFVQKMPGOLPILEGEVDSVKPGLSAFVDPQKGAETVQGLLEVAKDSIPRSHW

120

Db

61

RIHVYTFVQKMPGOLPILEGEVDSVKPGLSAFVDPQKGAETVQGLLEVAKDSIPRSHW

120

Qy

121

KKTPVVLKATAGLRLLPEHKAALLFEVKEIFPKSPFLVPKGSVINDGDEGLAWVTV

180

Db

121

KKTPVVLKATAGLRLLPEHKAALLFEVKEIFPKSPFLVPKGSVINDGDEGLAWVTV

180

Qy

181

NFLTQGLHGHROETVGTLDLGASTQITFLPQFEXTLEQTPRGYLTSEFMFNSTYKLYTH

240

Db

181

NFLTQGLHGHROETVGTLDLGASTQITFLPQFEXTLEQTPRGYLTSEFMFNSTYKLYTH

240

QY 241 SYLGFLKAARLALATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300
DB 241 SYLGFLKAARLALATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300
QY 301 EPCYAEVLVRVGRKGLHQPVEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
DB 301 EPCYAEVLVRVGRKGLHQPVEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
QY 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQLTKKVNNIETGVALGATFHL 420
DB 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQLTKKVNNIETGVALGATFHL 420
QY 421 LQSLGISH 428
DB 421 LQSLGISH 428

RESULT 4

US-09-240-639-9

; Sequence 9, Application US/09240639

; Patent No. 6350447

; GENERAL INFORMATION:

; APPLICANT: Chadwick, Brian Paul

; APPLICANT: Frischau, Anna-Maria

; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO CD39-LIKE

; FILE REFERENCE: POLYPEPTIDES AND NUCLEIC ACIDS

; CURRENT APPLICATION NUMBER: US/09/240,639

; CURRENT FILING DATE: 1998-01-29

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: Patentin ver. 2.0

; SEQ ID NO 9

; LENGTH: 428

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-240-639-9

Query Match 100.0%; Score 2250; DB 4; Length 428;
Best Local Similarity 100.0%; Pred. No. 6.4e-248;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATSWGTVPFMLVVSVCVSAVSHRNQOTWFBGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
DB 1 MATSWGTVPFMLVVSVCVSAVSHRNQOTWFBGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
QY 61 RIHYVTFVQKMPGQLPILGEVFDVSVKPGLSAPVDQPKQGAETVQGLLEVAKDSIPRSHW 120
DB 61 RIHYVTFVQKMPGQLPILGEVFDVSVKPGLSAPVDQPKQGAETVQGLLEVAKDSIPRSHW 120
QY 121 KKTPEVLKATAGLRLPEHKAKALLFEVKEIFRKSPLVPKGSVSINDGEGTAAWTV 180
DB 121 KKTPEVLKATAGLRLPEHKAKALLFEVKEIFRKSPLVPKGSVSINDGEGTAAWTV 180
QY 181 NFLTGQLHGHROETVGTDLGGASTQITFLPQFEKTLQTPRGYLTSPFNFSTYKLYTH 240
DB 181 NFLTGQLHGHROETVGTDLGGASTQITFLPQFEKTLQTPRGYLTSPFNFSTYKLYTH 240
QY 241 SYLGFLKAARLALATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300
DB 241 SYLGFLKAARLALATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300
QY 301 EPCYAEVLVRVGRKGLHQPVEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
DB 301 EPCYAEVLVRVGRKGLHQPVEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
QY 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQLTKKVNNIETGVALGATFHL 420
DB 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQLTKKVNNIETGVALGATFHL 420
QY 421 LQSLGISH 428
DB 421 LQSLGISH 428

RESULT 6

US-09-350-836B-5

; Sequence 5, Application US/09350836B

; Patent No. 6387645

; GENERAL INFORMATION:

; APPLICANT: Ford, John

; APPLICANT: Mulero, Julio

RESULT 5

US-09-350-836B-3

; Sequence 3, Application US/09350836B

; Patent No. 6387645

; GENERAL INFORMATION:

; APPLICANT: Ford, John

; APPLICANT: Mulero, Julio

; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE

; FILE REFERENCE: POLYPEPTIDES

; FILE REFERENCE: 28110/35761

; CURRENT APPLICATION NUMBER: US/09/350,836B

; CURRENT FILING DATE: 1999-07-09

; PRIOR APPLICATION NUMBER: 09/273,447

; PRIOR FILING DATE: 1999-03-19

; PRIOR APPLICATION NUMBER: 09/118,205

; PRIOR FILING DATE: 1998-07-16

; PRIOR APPLICATION NUMBER: 09/122,449

; PRIOR FILING DATE: 1998-07-24

; PRIOR APPLICATION NUMBER: 09/244,444

; PRIOR FILING DATE: 1999-02-04

; NUMBER OF SEQ ID NOS: 23

; SOFTWARE: Patentin ver. 2.0

; SEQ ID NO 3

; LENGTH: 428

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-350-836B-3

Query Match 100.0%; Score 2250; DB 4; Length 428;

Best Local Similarity 100.0%; Pred. No. 6.4e-248;

Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATSWGTVPFMLVVSVCVSAVSHRNQOTWFBGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
DB 1 MATSWGTVPFMLVVSVCVSAVSHRNQOTWFBGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
QY 61 RIHYVTFVQKMPGQLPILGEVFDVSVKPGLSAPVDQPKQGAETVQGLLEVAKDSIPRSHW 120
DB 61 RIHYVTFVQKMPGQLPILGEVFDVSVKPGLSAPVDQPKQGAETVQGLLEVAKDSIPRSHW 120
QY 121 KKTPEVLKATAGLRLPEHKAKALLFEVKEIFRKSPLVPKGSVSINDGEGTAAWTV 180
DB 121 KKTPEVLKATAGLRLPEHKAKALLFEVKEIFRKSPLVPKGSVSINDGEGTAAWTV 180
QY 181 NFLTGQLHGHROETVGTDLGGASTQITFLPQFEKTLQTPRGYLTSPFNFSTYKLYTH 240
DB 181 NFLTGQLHGHROETVGTDLGGASTQITFLPQFEKTLQTPRGYLTSPFNFSTYKLYTH 240
QY 241 SYLGFLKAARLALATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300
DB 241 SYLGFLKAARLALATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300
QY 301 EPCYAEVLVRVGRKGLHQPVEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
DB 301 EPCYAEVLVRVGRKGLHQPVEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
QY 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQLTKKVNNIETGVALGATFHL 420
DB 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQLTKKVNNIETGVALGATFHL 420
QY 421 LQSLGISH 428
DB 421 LQSLGISH 428

; EARLIER FILING DATE: 1999-02-04
; EARLIER APPLICATION NUMBER: 09/122,449
; EARLIER FILING DATE: 1998-07-24
; EARLIER APPLICATION NUMBER: 09/118,205
; EARLIER FILING DATE: 1998-07-16
; EARLIER APPLICATION NUMBER: 09/118,205
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-370-265-5

Query Match 100.0%; Score 2250; DB 4; Length 428;
Best Local Similarity 100.0%; Pred. No. 6.4e-248; Indels 0; Gaps 0;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATSWGTVFPMVLVVCVCSAVSHRNQQTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT 60
DB 1 MATSWGTVFPMVLVVCVCSAVSHRNQQTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT 60
QY 61 RIHVYTFVQKMPGOLPILLEGVDSVKPGLSAFVDOPKQGAETVOGLLEVAKDSIPRSHW 120
DB 61 RIHVYTFVQKMPGOLPILLEGVDSVKPGLSAFVDOPKQGAETVOGLLEVAKDSIPRSHW 120
QY 121 KKTVPVLKATAGLRLLPEHKAKALLFEVKEI FRKSPFLVPKGSVSI NDSDGGLAWTV 180
DB 121 KKTVPVLKATAGLRLLPEHKAKALLFEVKEI FRKSPFLVPKGSVSI NDSDGGLAWTV 180
QY 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSPFENSTYKLYTH 240
DB 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSPFENSTYKLYTH 240
QY 241 SYLGFGKLAARLATLGALETGDTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGF 300
DB 241 SYLGFGKLAARLATLGALETGDTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGF 300
QY 301 EPCYAEVLVRVGRKGLHOPEVQSGSFYAFSYTDDRAVDTMDIDYKGGILKVEDPERKAR 360
DB 301 EPCYAEVLVRVGRKGLHOPEVQSGSFYAFSYTDDRAVDTMDIDYKGGILKVEDPERKAR 360
QY 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQTKKVNNIETGVALGATFHL 420
DB 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQTKKVNNIETGVALGATFHL 420
QY 421 LQSLGISH 428
DB 421 LQSLGISH 428

RESULT 9

US-09-557-800C-3
; Sequence 3, Application US/09557800C
; Patent No. 6476211
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 28110/36457
; CURRENT APPLICATION NUMBER: US/09/557,800C
; CURRENT FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273447
; PRIOR FILING DATE: 1999-03-19

; PRIOR APPLICATION NUMBER: 09/122449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-557-800C-3

Query Match 100.0%; Score 2250; DB 4; Length 428;
Best Local Similarity 100.0%; Pred. No. 6.4e-248; Indels 0; Gaps 0;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATSWGTVFPMVLVVCVCSAVSHRNQQTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT 60
DB 1 MATSWGTVFPMVLVVCVCSAVSHRNQQTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT 60
QY 61 RIHVYTFVQKMPGOLPILLEGVDSVKPGLSAFVDOPKQGAETVOGLLEVAKDSIPRSHW 120
DB 61 RIHVYTFVQKMPGOLPILLEGVDSVKPGLSAFVDOPKQGAETVOGLLEVAKDSIPRSHW 120
QY 121 KKTVPVLKATAGLRLLPEHKAKALLFEVKEI FRKSPFLVPKGSVSI NDSDGGLAWTV 180
DB 121 KKTVPVLKATAGLRLLPEHKAKALLFEVKEI FRKSPFLVPKGSVSI NDSDGGLAWTV 180
QY 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSPFENSTYKLYTH 240
DB 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSPFENSTYKLYTH 240
QY 241 SYLGFGKLAARLATLGALETGDTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGF 300
DB 241 SYLGFGKLAARLATLGALETGDTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGF 300
QY 301 EPCYAEVLVRVGRKGLHOPEVQSGSFYAFSYTDDRAVDTMDIDYKGGILKVEDPERKAR 360
DB 301 EPCYAEVLVRVGRKGLHOPEVQSGSFYAFSYTDDRAVDTMDIDYKGGILKVEDPERKAR 360
QY 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQTKKVNNIETGVALGATFHL 420
DB 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQTKKVNNIETGVALGATFHL 420
QY 421 LQSLGISH 428
DB 421 LQSLGISH 428

RESULT 10

US-09-557-800C-5
; Sequence 5, Application US/09557800C
; Patent No. 6476211
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 28110/36457
; CURRENT APPLICATION NUMBER: US/09/557,800C
; CURRENT FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273447

; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/122449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-557-800C-5

Query Match 100.0%; Score 2250; DB 4; Length 428;
Best Local Similarity 100.0%; Pred. No. 6.4e-248;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATSGTVPFVFLVSCVSAVSHRNQQTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSGTVPFVFLVSCVSAVSHRNQQTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy 61 RIHYVTFVQKMPGQQLPILGEVFDVSKPCLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
Db 61 RIHYVTFVQKMPGQQLPILGEVFDVSKPCLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy 121 KKTVPVLKATAGLRLLPEHKAKALLFEVKIIFRKSPPFLVPKGSVSIIMDSDEGILAWVTY 180
Db 121 KKTVPVLKATAGLRLLPEHKAKALLFEVKIIFRKSPPFLVPKGSVSIIMDSDEGILAWVTY 180

Qy 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQFEKTLQTPRGYLTSEFNPSTYKLYTH 240
Db 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQFEKTLQTPRGYLTSEFNPSTYKLYTH 240

Qy 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVQYQYGGNQEVEGF 300
Db 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVQYQYGGNQEVEGF 300

Qy 301 EPCYAEVLVRVGRKQHPEEVQVGSFYAFSYVDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 301 EPCYAEVLVRVGRKQHPEEVQVGSFYAFSYVDRAVDTMDIDYKGGILKVEDFERKAR 360

Qy 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTKKVNNIETGALGATPHL 420
Db 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTKKVNNIETGALGATPHL 420

Qy 421 LQSLGISH 428
Db 421 LQSLGISH 428

RESULT 11
US-09-370-625A-3
; Sequence 3, Application US/09370625A
; Patent No. 6600032
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES
; FILE REFERENCE: 28110/35908
; CURRENT APPLICATION NUMBER: US/09/370,625A
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 428

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-370-625A-3

Query Match 100.0%; Score 2250; DB 4; Length 428;
Best Local Similarity 100.0%; Pred. No. 6.4e-248;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATSGTVPFVFLVSCVSAVSHRNQQTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSGTVPFVFLVSCVSAVSHRNQQTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy 61 RIHYVTFVQKMPGQQLPILGEVFDVSKPCLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
Db 61 RIHYVTFVQKMPGQQLPILGEVFDVSKPCLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy 121 KKTVPVLKATAGLRLLPEHKAKALLFEVKIIFRKSPPFLVPKGSVSIIMDSDEGILAWVTY 180
Db 121 KKTVPVLKATAGLRLLPEHKAKALLFEVKIIFRKSPPFLVPKGSVSIIMDSDEGILAWVTY 180

Qy 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQFEKTLQTPRGYLTSEFNPSTYKLYTH 240
Db 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQFEKTLQTPRGYLTSEFNPSTYKLYTH 240

Qy 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVQYQYGGNQEVEGF 300
Db 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVQYQYGGNQEVEGF 300

Qy 301 EPCYAEVLVRVGRKQHPEEVQVGSFYAFSYVDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 301 EPCYAEVLVRVGRKQHPEEVQVGSFYAFSYVDRAVDTMDIDYKGGILKVEDFERKAR 360

Qy 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTKKVNNIETGALGATPHL 420
Db 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTKKVNNIETGALGATPHL 420

Qy 421 LQSLGISH 428
Db 421 LQSLGISH 428

RESULT 12
US-09-370-625A-5
; Sequence 5, Application US/09370625A
; Patent No. 6600032
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDE
; FILE REFERENCE: 28110/35908
; CURRENT APPLICATION NUMBER: US/09/370,625A
; CURRENT FILING DATE: 1999-08-09
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-370-625A-5

Query Match 100.0%; Score 2250; DB 4; Length 428;
Best Local Similarity 100.0%; Pred. No. 6.4e-248;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATSGTVPFVFLVSCVSAVSHRNQQTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSGTVPFVFLVSCVSAVSHRNQQTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

QY 61 RIHVYTFVKMPGQLPILGEVFDSPKGLSAPVDPKQGAETVQGLLEVAKDSIPRSHW 120
DB 61 RIHVYTFVKMPGQLPILGEVFDSPKGLSAPVDPKQGAETVQGLLEVAKDSIPRSHW 120
QY 121 KKTVPVVKATAGRLRLPEHAKALLFEVKEIFRKSFPFLVPKGSVSIIMDSDEGILAWTV 180
DB 121 KKTVPVVKATAGRLRLPEHAKALLFEVKEIFRKSFPFLVPKGSVSIIMDSDEGILAWTV 180
QY 181 NFLTGQHLGHROETVGTDLGGASTQITFLPQFKTELEOTPRGYLTSPFNFSTYKLYTH 240
DB 181 NFLTGQHLGHROETVGTDLGGASTQITFLPQFKTELEOTPRGYLTSPFNFSTYKLYTH 240
QY 241 SYLGFLKKAARLALIGALETEGDTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGF 300
DB 241 SYLGFLKKAARLALIGALETEGDTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGF 300
QY 301 EPCYAEVLVRVGRGLHQPVEEVQSGFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
DB 301 EPCYAEVLVRVGRGLHQPVEEVQSGFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
QY 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQTKKNNIETGVALGATPHL 420
DB 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQTKKNNIETGVALGATPHL 420
QY 421 LQSLGISH 428
DB 421 LQSLGISH 428

RESULT 13
US-09-608-285A-7
; Sequence 7, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
; FILE REFERENCE: 28110/36570
; CURRENT APPLICATION NUMBER: US/09/608,285A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/583,231
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1998-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1998-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-608-285A-7

method of antibody

Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 MATSWGTVFFMLVWSCVSAVSHRNQQTWFEIGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
DB 1 MATSWGTVFFMLVWSCVSAVSHRNQQTWFEIGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
QY 61 RIHVYTFVKMPGQLPILGEVFDSPKGLSAPVDPKQGAETVQGLLEVAKDSIPRSHW 120
DB 61 RIHVYTFVKMPGQLPILGEVFDSPKGLSAPVDPKQGAETVQGLLEVAKDSIPRSHW 120
QY 121 KKTVPVVKATAGRLRLPEHAKALLFEVKEIFRKSFPFLVPKGSVSIIMDSDEGILAWTV 180
DB 121 KKTVPVVKATAGRLRLPEHAKALLFEVKEIFRKSFPFLVPKGSVSIIMDSDEGILAWTV 180
QY 181 NFLTGQHLGHROETVGTDLGGASTQITFLPQFKTELEOTPRGYLTSPFNFSTYKLYTH 240
DB 181 NFLTGQHLGHROETVGTDLGGASTQITFLPQFKTELEOTPRGYLTSPFNFSTYKLYTH 240
QY 241 SYLGFLKKAARLALIGALETEGDTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGF 300
DB 241 SYLGFLKKAARLALIGALETEGDTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGF 300
QY 301 EPCYAEVLVRVGRGLHQPVEEVQSGFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
DB 301 EPCYAEVLVRVGRGLHQPVEEVQSGFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
QY 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQTKKNNIETGVALGATPHL 420
DB 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQTKKNNIETGVALGATPHL 420
QY 421 LQSLGISH 428
DB 421 LQSLGISH 428

RESULT 14
US-09-350-836B-7
; Sequence 7, Application US/09350836B
; Patent No. 6387645
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; FILE REFERENCE: 28110/35761
; CURRENT APPLICATION NUMBER: US/09/350,836B
; CURRENT FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-836B-7

method of ID inhibition

Query Match 99.3%; Score 2235; DB 4; Length 428;
Best Local Similarity 99.3%; Pred. No. 3.3e-246;
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 MATSWGTVFFMLVWSCVSAVSHRNQQTWFEIGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
DB 1 MATSWGTVFFMLVWSCVSAVSHRNQQTWFEIGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
QY 61 RIHVYTFVKMPGQLPILGEVFDSPKGLSAPVDPKQGAETVQGLLEVAKDSIPRSHW 120

Query Match 99.3%; Score 2235; DB 4; Length 428;
Best Local Similarity 99.3%; Pred. No. 3.3e-246;

Db 61 RIHVYFVQKMPQGPILLEGVFDVSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
Qy 121 KKTPTVVKATAGLRLPEHKAALIFEVKEIFRKSPLVPKGSVIMTQDEGIFAWVTY 180
Db 121 KKTPTVVKATAGLRLPEHKAALIFEVKEIFRKSPLVPKGSVIMTQDEGIFAWVTY 180
Qy 181 NPLTQGLHGHQETVGTDLGGASTQITFLPQPEKTLBQTPRGYLTSPFNSTYKLYTH 240
Db 181 NPLTQGLHGHQETVGTDLGGASTQITFLPQPEKTLBQTPRGYLTSPFNSTYKLYTH 240
Qy 241 SYLGFGKAAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKKYQYGGNQEVEVGF 300
Db 241 SYLGFGKAAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKKYQYGGNQEVEVGF 300
Qy 301 EPCYAEVLVRVVRGKLHQPVEVQSGSFYAFSYYYDRAVDTDMDIYKGGILKVEDFERKAR 360
Db 301 EPCYAEVLVRVVRGKLHQPVEVQSGSFYAFSYYYDRAVDTDMDIYKGGILKVEDFERKAR 360
Qy 361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGALGATPHL 420
Db 361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGALGATPHL 420
Qy 421 LQSLGISH 428
Db 421 LQSLGISH 428

RESULT 15

US-09-370-265-7
Sequence, Application US/09370265
Patent No. 6447771
GENERAL INFORMATION

APPLICANT: Eads, John
APPLICANT: Mulero, Julio
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE

FILE OF INVENTION: POLYPEPTIDES
FILE REFERENCE: 28111/35908
CURRENT APPLICATION NUMBER: US/09/370,265

EARLIER FILING DATE: 1999-08-09
EARLIER FILING DATE: 1999-07-16
EARLIER FILING DATE: 1999-07-16

EARLIER FILING DATE: 1999-07-09
EARLIER FILING DATE: 1999-03-19
EARLIER FILING DATE: 1999-03-19

EARLIER FILING DATE: 1999-02-04
EARLIER FILING DATE: 1999-02-04
EARLIER FILING DATE: 1998-07-24

EARLIER FILING DATE: 1998-07-16
NUMBER OF SEQ ID NOS: 37
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 7
LENGTH: 428
TYPE: PRI

ORGANISM: Homo sapiens
US-09-370-265-7

Query Match 99.3%; Score 2235; DB 4; Length 428;
Best Local Similarity 99.3%; Pred. No. 3.3e-246;
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MATSWGTVPFMLVVCVCSAVSHRNQOTWFGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSWGTVPFMLVVCVCSAVSHRNQOTWFGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Qy 61 RIHVYFVQKMPQGPILLEGVFDVSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
Db 61 RIHVYFVQKMPQGPILLEGVFDVSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
Qy 121 KKTPTVVKATAGLRLPEHKAALIFEVKEIFRKSPLVPKGSVIMTQDEGIFAWVTY 180

Search completed: July 1, 2004, 13:50:10
Job time : 17.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2004, 13:43:42 ; Search time 16.5 Seconds

(without alignments)
2495.150 Million cell updates/sec

Title: US-10-091-085-3

Perfect score: 2250

Sequence: 1 MATSWGTVFPMVVCVCSA.....ETGVALGATPHLLQSLGISH 428

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: PIR.78.*
- 2: PIR1.*
- 3: PIR3.*
- 4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	616.5	27.4	479	2 T23508	hypothetical prote
2	538.5	23.9	556	2 T39109	probable guanosine
3	489	21.7	455	2 S48859	nucleoside triphos
4	483.5	21.5	518	2 A40732	guanosine-diphosph
5	446.5	19.8	454	2 J04616	aprase (EC 3.6.1.
6	425.5	18.9	516	2 G84442	probable nucleosid
7	397	17.6	483	2 D86276	hypothetical prote
8	387.5	17.2	485	2 T34147	hypothetical prote
9	372.5	16.6	557	2 T16696	hypothetical prote
10	372	16.5	630	2 S50463	hypothetical prote
11	351	15.6	510	2 I56242	lymphoid cell acti
12	347	15.4	572	2 T40856	probable nucleotid
13	329.5	14.6	405	2 S86276	hypothetical prote
14	274	12.2	1052	2 T04439	hypothetical prote
15	248	11.0	508	2 C86276	7A19.33 protein -
16	138	6.1	628	2 A55421	nucleoside-triphos
17	110	4.9	497	1 J02192	subtilisin-like pr
18	106.5	4.7	553	1 G1BPSV	gene 1 protein - s
19	105.5	4.7	774	2 T14555	DNA polymerase hom
20	104.5	4.6	684	2 T47694	probable serine/th
21	104.5	4.6	4427	2 P06037	polyketide synthas
22	102.5	4.6	1019	2 T40813	probable cell divi
23	101.5	4.5	535	2 P97910	glucan 1,6-alpha-g
24	101.5	4.5	1151	2 H71347	hypothetical prote
25	99.5	4.4	535	2 C95040	glucan 1,6-alpha-g
26	99.5	4.4	548	1 HYBSS	bacillolysin (EC 3
27	97.5	4.3	604	2 F89453	protein F35H12.4 [
28	97.5	4.3	654	2 AG3522	iron-regulated out
29	97	4.3	1220	2 T18291	patched protein -

30	96	4.3	711	2 S66749	hypothetical prote
31	95.5	4.2	2076	2 S15999	fatty-acyl-CoA syn
32	95	4.2	652	1 JC2191	subtilisin-like pr
33	93	4.1	478	2 P90497	hypothetical prote
34	93	4.1	635	1 T37835	probable phosphos
35	93	4.1	962	2 JC5571	subtilisin-like pr
36	93	4.1	969	1 A39490	subtilisin-like pr
37	93	4.1	975	2 JC5570	coproporphyrinogen
38	92	4.1	290	2 A87312	H+-transporting tw
39	92	4.1	528	2 S13641	glycine hydroxymet
40	91	4.0	417	2 H70174	t-plasminogen acti
41	91	4.0	477	2 J80597	glucanase kinase g
42	91	4.0	500	2 D83984	probable ferric si
43	91	4.0	726	2 AB0122	phosphatidylserine
44	90.5	4.0	301	2 E71482	succinyl-diaminopi
45	90.5	4.0	377	2 C82113	

ALIGNMENTS

RESULT 1

T23508

hypothetical protein K08H10.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000

C:Accession: T23508

R:Gardner, A.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19750

A:Accession: T23508

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-479 <WIL>

A:Cross-references: EMBL:Z83113; PIDN:CAB05544.1; GSPDB:GNO0023; CESP:K08H10.4

A:Experimental source: clone K08H10

C:Genetics:

A:Gene: CESP:K08H10.4

A:Map position: 5

A:Introns: 36/3; 83/3; 189/1; 300/2; 412/3

C:Superfamily: nucleoside triphosphatase chromatin-associated

Query Match	27.4%	Score	616.5	DB	2	Length	479
Best Local Similarity	34.3%	Pred. No.	1.3e-43				
Matches	150	Conservative	81	Mismatches	167	Indels	39
Gaps	16						
Qy	9	FMLVVS--CVCASVSHRNOQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGRIHYVT	66				
Db	6	FSILLISFSLSSVVTTKQY-WCHGGVILN---NQHTCRFTTIVIDAGSTGTRLLYK	60				
Qy	67	FVQK-----MPQLPILEGEVDSVKPGLSAFVPOPKQCAETVQGLLEVAKDSIPRSH	119				
Db	61	FIDHPAIAHGMPFK---VEKEIFQVKKPGLSSFAKSPSSAADSLEPLLQARKEVPHEM	117				
Qy	120	WKTPVVLKATAGRLRLPEHKAKALLFEVKE-IFRKSPPFLVPGKSVISIMGDSBGLIAWV	178				
Db	118	WEXTPTLTKATAGRLRLPGDMADDILESVEERIFNSGFFAAPPDAVNVMPGSDGVSWF	177				
Qy	179	TNPLATQLH-----GHR---QETVGLDGLGASTQITPLPOFEXTLEOTPGYLTSFE	229				
Db	178	TLNILETFTDPTVGHKFAHRSVAAPDLGGSTQLTWPNNEAVFSHV-GYERDID	236				
Qy	230	MFNSTYKLYTHSYGLFGLKAARLATIQALETG--TDGHTFRACAPRWLE-AEWIPGVV	286				
Db	237	PFCHIRLPTHTSPGLNGLIAARLNIL-QLETDNEIESTHOLITSCEMPEGYQLTEWEY-AL	294				
Qy	287	KYQYGNQGEVCGVEFCYAEVLVWR-GKLHQEEVQVGSFYAFSYVYDRAVDTMDIYE	345				
Db	295	KF--WNINGSSSHSPFCYGTGTTKNFVSESSEIMHURELKGSPVLFSTFFDRLALNSGLVKN	353				
Qy	346	KGKILKVEDEFERKAREVC-----DNLENFTSGSPFLCMLDSYITALLKQGFADSTVQL	401				
Db	354	EGGKIELRQFKEAAEIAACREKTEIDDGSHWMPQCLDLTYIVSLLRDGYQFEDNQPLVL	413				

QY 402 TKXVNIETGHALGATE 418
Db 414 AKKIKMEVSWCQGLAF 430

RESULT 2
T39109
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T39109
R:Sattell, B.G.; Rajandream, M.A.; Quail, M.; Seegar, K.; Harris, D.
submitted to the EMBL Data Library, October 1999
A:Reference number: Z21828
A:Accession: T39109
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-356 <BAR>
A:Cross-references: EMBL:AL121741; PIDN:CA857338.1; GSPDB:GN00066; SPDB:SPAC824.08
A:Experimental source: strain 972h-; cosmid c824
C:Genetics:
A:Gene: SPDB:SPAC824.08
A:Map position: 1

Query Match 23.9%; Score 538.5; DB 2; Length 556;
Best Local Similarity 34.4%; Pred. No. 5.5e-37;
Matches 144; Conservative 61; Mismatches 157; Indels 57; Gaps 13;

QY 49 YGIMFDAGSTGTRIHYVTFVQKMPQLPILEGEVDSVKPGLSAFVQDPKQGAETVQGL 108
Db 134 YVLMIDASTGSRVHYQFNCPNPS--PKLEEFPWTEPGLSSPAGPPEGANASLDPL 191
QY 109 EVAKDSIPRSHWKTPVVLKATAGLRLPRHKAKALLFEVKEIFPKRS--PFLVPKXGSVIM 167
Db 192 DYAMENVPEEVRKCPPIAVKATAGLRLGCESEKAILKSVRQHLENDYPPFLVXGVGSL 251
QY 168 DGSDEGILAWTVNFLTQGLHGH--RQETVGLDGLGASTQITFLPQPEKTELEQTPRG--- 223
Db 252 ESMGEGIVAMITINLLGTGKATHSTVAVMDLGGASTQVLFPRFASDGSILVDGDKH 311
QY 224 YLTSEFMFNSTYKLYTHSYLGFGLKAARL-----ATLGALETEGTDGHTFRSAC 272
Db 312 YVLDYN--GEQVELYQHSYGLKAEARKLHKFVLNNAEALKESSLELLG-DSTSIIHP 368
QY 273 L-----PRWEAEWIFGVKYQYGGNQEVEGPEPCYAEVLVRV-----GKLHOP 318
Db 369 LHLNASTHPDSKSEASEVVFVGPSELALHLSQCRGIAEKALYKXKNCVPRPCSFNGVHOP 428
QY 319 ---EBVQKGSFYAFSYVYDRAVDTDMDIYKGGILKVEDFERKAREVC----- 363
Db 429 KTEPTDTPSIVLISYFYDR-----WISLGMPTFTIEDMKYLANSVCSGGTYWQDAFSL 483
QY 364 -DNLNFTGSGSFLCMDSYITALLKDGFGFADSTVLQTKKNNIETGNALGATPHLL 421
Db 484 TDALKELKE-EPEWCLDNLNLSLSVGVYEIPNNRQLHTAKKIDNKELGWCIGASLSML 541

RESULT 3
S48859
nucleoside triphosphatase precursor, chromatin-associated - garden pea
C:Species: Pisum sativum (garden pea)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 29-Sep-1999
C:Accession: S65147; S48859
R:Hsieh, H.L.; Tong, C.G.; Thomas, C.; Roux, S.J.
Plant Mol. Biol. 30, 135-147, 1996
A:Title: Light-modulated abundance of an mRNA encoding a calmodulin-regulated, chromatin
A:Reference number: S65147; M01D:96197404; PMID:8616230
A:Accession: S65147
A:Molecule type: mRNA
A:Residues: 1-455 <HS2>
A:Cross-references: EMBL:Z32743; NID:G563611; PIDN:CAA83655.1; PID:G563612
C:Superfamily: nucleoside triphosphatase chromatin-associated

C;Keywords: nucleus

Query Match 21.7%; Score 489; DB 2; Length 455;
Best Local Similarity 33.7%; Pred. No. 5.8e-33;
Matches 140; Conservative 62; Mismatches 160; Indels 54; Gaps 14;

QY 49 YGIMFDAGSTGTRIHYVTFVQKMPQLPILEGEVDSVKPGLSAFVQDPKQGAETVQGL 107
Db 44 YAVVEDAGSTGSRHVVHFHQNLDLLHIGKGVYNNKITPGLSSVANNPEQAASLIPL 102
QY 108 LEVAKDSIPRSHWKTPVVLKATAGLRLPRHKAKALLFEVKEIFPKRSFLVPKXGSVSI 166
Db 103 LEQAEVVPDDLOKFTPVRLGATAGLRLGDSASEKILQSVRMDLNSRSTFNQPDVAISI 162
QY 167 MDGSEGILAWTVNFLTQGLHGHROETVGLDGLGASTQITFLPQPEKTELEQTPR--- 222
Db 163 IDGTQEGSYLAVTVNVALGNLKKYTKTVGVLDLGGSGVQMAVAVS-KTKAKNPKVADG 221
QY 223 --GYLTSEFMFNSTYKLYTHSYLGFGLKAARLALATLGALETEGTDGHTFRS--ACLPKWL 278
Db 222 DDPYIKKVLKIGIPYDLYVHSHYLFQREASRAILKL-----TPRSPNCPCLLAGPN 272
QY 279 AEWIEGKVKYQYGGNQEVEGPEPCYAEVLVRVVGKLGHOPEEVQKRSF----- 326
Db 273 GIYITSGEFPKATAYTSG--ANFNKCKNTIRKAL--KLNYPCTQNCFTGGIWWGGGNGQ 329
QY 327 ---YAFSYVYDRAVDTDMDIYKGG--ILKYVEDFERKAREVCD-NLENFTSGSPFL----- 376
Db 330 KNLFPASSSPFLPDTGMDVASTPNTFLRPVDIETRAKEACALNFEDAKSTYFPFLDKKNV 389
QY 377 ---CMLDSYITALLKDGFGFADSTVLQTKKNN-----LETGALGATPHLLQSL 424
Db 390 ASYVCMDDLIYQVLLVDGFLDPLQKITSGKBIYQDAIVEAAWPLGNVAEISAL 445

RESULT 4
A40732
guanosine-diphosphatase (EC 3.6.1.42) - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YEL042w
C:Species: Saccharomyces cerevisiae
C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 02-Sep-2000
C:Accession: A40732; B40732; S30837; S50502
R:Abelton, C.; Yanagisawa, K.; Mandon, E.C.; Hausler, A.; Moremen, K.; Hirschberg, C.
J. Cell Biol. 122, 307-323, 1993
A:Title: Guanosine diphosphatase is required for protein and sphingolipid glycosylat
A:Reference number: A40732; M01D:93308137; PMID:8391537
A:Accession: A40732
A:Molecule type: DNA
A:Residues: 1-518 <ABE>
A:Cross-references: EMBL:L19560; NID:G349392; PIDN:AAA34656.1; PID:G349393
A:Note: sequence extracted from NCBI backbone (NCBIN:134708, NCBIP:134711)
A:Accession: B40732
A:Molecule type: protein
A:Residues: 125-144; 238-257; 276-281; 366-374; 399-412 <AB2>
R:Mulligan, J.T.; Dietrich, F.S.; Hennessey, K.M.; Sehl, P.; Komp, C.; Wei, Y.; Tayl
submitted to the EMBL Data Library, February 1993
A:Reference number: S30812
A:Accession: S30837
A:Molecule type: DNA
A:Residues: 1-518 <MUL>
A:Cross-references: GB:U18779; EMBL:U10830; NID:G603625; PIDN:AA865000.1; PID:G60363
R:Dietrich, F.S.
A:Description: The sequence of S. cerevisiae cosmid 8199, 8334, and 9871.
A:Reference number: S50491
A:Accession: S50502
A:Molecule type: DNA
A:Residues: 1-518 <DIE>
A:Cross-references: EMBL:U18779; NID:G603625; PIDN:AA865000.1; PID:G603637; MIPS:YEL
C:Genetics:
A:Gene: SGD:GDAL
A:Cross-references: SGD:S0000768; MIPS:YEL042w
A:Map position: 5L

C;Function:
A;Description: hydrolase
C;Superfamily: nucleoside triphosphatase chromatin-associated
C;Keywords: Glycoprotein; Golgi apparatus; hydrolase; transmembrane protein
F;10-24/Domain: Transmembrane #status predicted <TM>
F;41,280,335/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 21.5%; Score 483.5; DB 2; Length 518;
Best Local Similarity 32.8%; Pred. No. 28-32;
Matches 142; Conservative 56; Mismatches 160; Indels 75; Gaps 15;
QY 49 YGIMFAGSTGTHVTF-VQXMPGQLPILEGVPSVXPKGLSAFVDPQKQAGTVOGL 107
DB 93 YVIMIDAGTGRVHIVKFDVCTSP---PTLDEKEDMLEPGLSSFDTSVGAANSLDPL 149
QY 108 LEVAKDISPRSHMKKTPVTLKATAGLRLPEHKAKALLFEVKIIFRKS-PFLVPGK-SVS 165
DB 150 LKVAMNVPIKARSCTPVAVKATAGLRLGDAKSKILSAVRDHLKDYPPVVEGDGVS 209
QY 166 IMDSGEGILAWTVNPLTQOL-HGHRQETVGTDLGGASTOITLPPQEKTELEQTPRG 223
DB 210 IMGDEBGFAMITNTLLGNMGANGPKLPTAAVFDLGGSTQIVPEPTPIKEKVDGE 269
QY 224 YLTSFEMFNSTYKLYTHSYLGFGLKAAR-----LATLGALETEGTDGTFERSAC 272
DB 270 HKPDLKFDENYTLQPSHLGYGLKEGRNVNSVLVENALKDQKILKGDNTKTHQLSSPC 329
QY 273 LPRWLEA-FWI-----FGVYQYGG-----NQGEVGFEPYAEVLV 309
DB 330 LPPKVNATNPKVTLSEKVTVIDFIDPEPSGAQRFLEILNKDAQSPSPCSF---- 385
QY 310 VWRGKLHQPBEV-----QRGSFYAFSYYYDRAVDTDM-IDYEGGILKVEDFEKAREVC- 363
DB 386 ---NGVHQPSLVTRFEKSDNIYIFSYFYDTRPLGMLSP-----TLNELNDLARI 436
QY 364 -----DNLNFTSGSPFLMDLSYITALLKQFGADSTVLOLTQKVNNIET 410
DB 437 GEETWNSVFGIAGLSELESDSHP-CLDLSFQVSLHTGYDIPLORELRTGKKIANKEI 495
QY 411 GWALGATFHLLOS 423
DB 496 GWCLGASLPLIKA 508

RESULT 5
JC4616
apyrase (EC 3.6.1.5) precursor - potato
X;Alternate names: adenylpyrophosphatase; ATP-diphosphohydrolase
C;Species: Solanum tuberosum (potato)
C;Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 28-May-1999
C;Accession: J04616; PC4147
R;Handa, M.; Guidotti, G.
Biochem. Biophys. Res. Commun. 218, 916-923, 1996
A;Title: Purification and cloning of a soluble ATP-diphosphohydrolase (Apyrase) from potato
A;Reference number: J04616; MUID:96158985; PMID:8579614
A;Accession: J04616
A;Molecule type: mRNA
A;Residues: 1-454 <HAN>
A;Cross-references: GB:U58597; MID:g1381632; PIDN:AAB02720.1; PID:g1381633
A;Accession: PC4147
A;Molecule type: protein
A;Residues: 59-95;96-131,132-160;236-253;332-345 <HA2>
A;Experimental source: tubers
A;Note: The authors translated the codon GCA for residue 215 as Gly
C;Comment: This enzyme belongs to a family of E-type ATPases, and it catalyzes the hydrolysis of ATP to ADP and inorganic phosphate, a reaction that is essential for the enzyme has nucleotide substrate specificity, divalent cation requirement, and insensitivity to starch synthesis.
C;Genetics:
A;Gene: xrop1
C;Superfamily: nucleoside triphosphatase chromatin-associated
C;Keywords: Glycoprotein; hydrolase; transmembrane protein
F;1-30/Domain: signal sequence #status predicted <SIG>
F;8-25/Domain: transmembrane #status predicted <TM>

F;11-454/Product: ATP-diphosphohydrolase #status predicted <MAT>
F;44-65/Region: actin-heat shock protein 70-hexokinase beta-phosphate binding
F;44-65/Region: nucleotide binding #status predicted
F;192-212/Region: actin-heat shock protein 70-hexokinase gamma-phosphate binding
F;192-212/Region: nucleotide binding #status predicted
F;390-410,427-446/Region: hydrophobic carboxyl end
F;151,262/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 19.8%; Score 446.5; DB 2; Length 454;
Best Local Similarity 28.8%; Pred. No. 2.1e-29;
Matches 131; Conservative 78; Mismatches 177; Indels 69; Gaps 16;
QY 25 NQOTWPEGIFLS-----SMCPINVA-----STLYGIMFAGSTGTHVTF 65
DB 3 NQNSHFIFILALFLVPLSLSKKNVNAQIPLRRULLSHSEHYAVIFDAGSTGSRVHF 62
QY 66 TVQXMPGQLPILEG-EVPSVXPKGLSAFVDPQKQAGTVOGLLEVAKDSIPRSHMKKTP 124
DB 63 REDEKL-GLLPTGNINIEYPMATEPGLSSYAEKPKAAANSLPDLGAEVVPORLOSETP 121
QY 125 VVLKATAGLRLPEHKAKALLFEVKIIFR-KSPFLVPGKSVSIMDGDGILAWTVNPL 183
DB 122 LELGATAGLRLPEHKAKALLFEVKIIFR-KSPFLVPGKSVSIMDGDGILAWTVNPL 181
QY 184 TQQLGHRQETVGTDLGGASTOITL-----POPEKTELE-OTPRGYLTSPFEMFNSTYKLYT 239
DB 182 LGNMGKDYESTTATIDLGSGSVQMAIYASNEQFAKAPQNEDEGEPYVQOQKELMSKDNLYV 241
QY 240 HSYVGLPKAARLATLGALETEGTDGTFERSACLPRLWLEAEIFGVYQYGGNQGEV 299
DB 242 HSYVGLPKAARLATLGALETEGTDGTFERSACLPRLWLEAEIFGVYQYGGNQGEV 292
QY 300 FEPYAEVLVVRG--KLHQPBEVQSGP-----YAFSYYYDRAVDTDMI 342
DB 293 ---SWKRCRLTRHALKINAKNIECTFNGVWGGGQKQKNIHASSPFYDGAQVIV 349
QY 343 DYE-KGILLKVEDFEKAREVC-DNLENFTS-----GSPFLMDLSYITALLKQFG 392
DB 350 DTKFPKALAKPQYLNAAKVAQCTNVADIKSIPPKTQDRNIPFLCMDLIYEYTLVDGFG 409
QY 393 F---ADSTVLOLTQKVN--IETGALGATFHLLOS 423
DB 410 LNPHEKITVHDVQYKMYLVGAAPLGLCAIDLVS 444

RESULT 6
G84442
probable nucleoside triphosphatase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 18-Nov-2002
C;Accession: G84442
R;Jin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Varakan, S.E.; Umayam, L.; Tallon
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Vente
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: G84442
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-516 <STO>
A;Cross-references: GB:AE002093; MID:g3461821; PIDN:AAC32915.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g02970
A;Map position: 2
C;Superfamily: nucleoside triphosphatase chromatin-associated
Query Match 18.9%; Score 425.5; DB 2; Length 516;
Best Local Similarity 30.6%; Pred. No. 1.5e-27;
Matches 127; Conservative 61; Mismatches 154; Indels 73; Gaps 13;
QY 42 NVSASTLYGIMFAGSTGTHVTFVQXMPGQLPILEGVPSVXPKGLSAFVDPQKQAGTVOGL 97

```
Db 61 NURGLRYSVVVDGSGTGTRIHVFGY--RIESGKPVFFRGANYASLKLHGLSAPADDP 118
Qy 98 KQAEVTVQGLLEVAKDSIPRSHWKTTPVVLKATAGRLILPEHAKALFEVKEIFRKSPP 157
Db 119 DGASVSLTELVEFAKGRVSKGMWIEVEVELMATAGXRLLELPVQEKILGVARRVLKSGF 178
Qy 158 LVPKGSVSMDSGRLAWTVNLTGQLHGHROETVGTLDLGGASTQITFL-----P 211
Db 179 LFRDEWASVISGDEGVAVWVANFALSGDGLKTTGIVELGASAGVTFVPSSEPP 238
Qy 212 QPEKLEQTPRGYLSFFENFSTYKLYTHSYLGFGLKAARLATLAL-----ETEGTD 264
Db 239 EFSRIT-----SFG--NVTNLYSHSPLHFGNAADKLMGSLLRDHSNAVEPTR 287
Qy 265 GHIFRSACLP-----WLEAWIFGGVYQYGGNQEGVEGPEPCYAVLWV 311
Db 288 EXIFDPCAPKGYNDANTQKLSGLLAEESLSDFSQAGN-----YSQCRSAALTIL 341
Qy 312 RGLKHQPEVQVQSGSFVAFSYVDRAVDTMI-----DYEGKGLLKVEDFERKAREV 362
Db 342 QDQNGRILLIIAGFSPLFGLGKAWLSNWSIAGERFCGEDWSK---LRVKDPSLHESDL 398
Qy 363 CONLENFTSGSPFLCMLSYITALLKDGFGP-ADSTVLQLTKKVNNIETGMALGA 416
Db 399 LR-----YCFSSAYIVSLHDTLGLDDEIRIGYANQAGDIPLDWALGA 442

RESULT 7
D86276
hypothetical protein F7A19.34 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 19-Nov-2002
C:Accession: D86276
R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maifi, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MID:21016719; PMID:11130712
A:Accession: D86276
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-483 <STO>
A:Cross-references: GB:AE005172; NID:G5080801; PIDN:AD39311.1; GSPDB:GNO0141
C:Genetics:
A:Map position: 1
C:Superfamily: nucleoside triphosphatase chromatin-associated

Query Match 17.6%; Score 397; DB 2; Length 483;
Best Local Similarity 27.6%; Pred. No. 3.3e-25;
Matches 120; Conservative 71; Mismatches 162; Indels 82; Gaps 17;

Qy 32 GIFLSSMCPINVSASTLYGIMFDAGSTGTRIHV--YTFVQKMPGQLPILEGEVDP----- 84
Db 59 GSLLSRCKLR-----YSLVIDAGSSGTRVHVFGYWFPSGKP-----VPDFGEKH 103
Qy 85 ----SVKPGLSAFVDPQKGAETVQGLLEVAKDSIPRSHWKTTPVVLKATAGRLILPEKH 140
Db 104 YANILKLTGGLSSVADNPEGASVSVTKLVEPAQRIPKRMFRSDIRLMATAGRLLEVPV 163
Qy 141 AKALFEVKEIFRKSFPVLPKGSVSMDSGRLAWTVNLTGQLHGHROETVGTLDL 200
Db 164 QEQLLEWTRVRLRSSGFMFRDEWANVISGDEGIYSWITANYALASLGLDPLETTGIVEL 223
Qy 201 GGASTQITLPOFEKTLQTPRGYLSFFENFSTYKLYTHSYLGFGLKAARLATGALET 260
Db 224 GGASQVTVF-----SSEHVPPEYSKTIAYGNISYTIYSHSFLDYKDAALKLEKLN 278
```

RESULT 8

T34147

hypothetical protein C33H5.14 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 17-Mar-2000

C:Accession: T34147

R:Bradshaw, H.; Stellyes, L.

submitted to the EMBL Data Library, November 1995

A:Description: The sequence of C. elegans cosmid C33H5.

A:Reference number: Z21482

A:Accession: T34147

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-485

A:Cross-references: EMBL:U41007; PIDN:AAA82272.1; CESP:C33H5.14

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:C33H5.14

A:Introns: 22/1; 83/1; 120/2; 167/1; 269/3; 399/3

C:Superfamily: nucleoside triphosphatase chromatin-associated

Query Match 17.2%; Score 387.5; DB 2; Length 485;

Best Local Similarity 27.7%; Pred. No. 2.1e-24;

Matches 120; Conservative 84; Mismatches 152; Indels 77; Gaps 18;

Qy 49 YGIMFDAGSTGTRIHVTVFQKMPGQLPILEGEVDS-----VKPGLSAFVDPQKGA 102

Db 25 YGVICDAGSSGTRULFVIT-LKPLSGGLTNIDTLHESEPVVKKVTPGLSSFGDKPEQVVE 83

Qy 103 TVQGLLEVAKDSIPRSHWKTTPVVLKATAGRLILPEHAKALFEVKEIFRKSFPFL-VPK 161

Db 84 YLTPILPFAEHHIPEYQLGETDLIFATAGWELLPEAQKDALIKNLQGLKSVTALRVSD 143

Qy 162 GSVSIMDSDEGILAWTVNLTGQLHGHROETVGTLDLGGASTQITFLPOFEK----- 215

Db 144 SNIRIIDGAMBGIIYSWIAVNVILGRFDKENDSKVGMIDMGASVQIAPEIANEKESYNGG 203

Qy 216 TLEQTPRGYLSFFENFSTYKLYTHSYLGFGLKAARLATGALETGDTGHTFRSACLPR 275

Db 204 NYVEINLGSITNEDYK--YKIYSTTFLGYGANGLKXENSLVKSGNS-----NDSCSPR 257

Qy 276 WLEAWIFGGVYQYGGNQEGVEGPEPCYAVLWVTVGKLHOPE----- 319

Db 258 GLNR--LIG-----EFTVNGTGE--WDVCLAQVSLI-GDKAQPSCPNPTCFLRNVAPSV 308

Qy 320 EVQSGSFVAFS-YYVDRAVDTMDIDYKGGITLKVEDFERKAREVC-----DNLENFTSGSP 374

Db 309 NLSTVQLXGSEYVWTTSS-----NFGSGGEYHVKFTDEVRYKYCKQKMDNDIQDFKENE 362

Qy 375 F-----LQMDLSYITALLKDGFGPADST--VLQTKUNNIETGMALGA----- 416

Db 363 FPNADIBRLGTCFKAANVTISVLHDGFN-VDTKHLPQSVLKIAGEEMQWALGAMLYHSK 421

Qy 417 --TFHLLQSLGIS 427

Db 422 DLKENLLEQUELEVA 434

Db 152 VERSLSNYPF--DFQGAIIITQEGAYGVITINVLKQSKQTRNFSIVPYETNNQET 209

Qy 196 GTLDGGASTQITLPOPEKTELETPRGYLTSEFNFSTYKLYTHSYLGFGLKAARLATL 255

Db 210 CALDGGASTQVTFVQ--NQIIE--SPDNAL--QPRLYGKQVNVYTHSYFLCYGKQALWQL 266

Qy 256 CALTEGTGHTFRSACL-----PAWLEAEWIFGKVKYQYQNGE 297

Db 267 -AKDIQVANSBILRDPCHPGYKVVNSDYKTPCTKRFMTLPFPQFBIQIGN---- 321

Qy 298 VGFPPCAVAELRV-----VRGKLQPEVEVQSGYAFSYYYDRAVDMDIDYE 345

Db 322 --YQCHQSILELNTSYCPYSQCAFNGLPPLQDGFASAF-YFVMKFLN---LTSE 375

Qy 346 KGGILKVEDFERK-AREVCDNLENFTSG--SPFL---CMDSYITALLKQDGF-ADS-T 397

Db 376 KVSQEKVTMMKKFCAQPEBIEKTSYAGVKEKYLSEYCFSGYITLSLLQGYHTADSWE 435

Qy 398 VLQLTKKVNNITETGALCATPHL 420

Db 436 HIHFGKIQSGDAGHTLGYMLNL 458

RESULT 12

T40856

probable nucleotide phosphatase - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T40856

R:Ramsperger, U.; Pohl, T.; Wood, V.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, October 1999

A:Reference number: Z21952

A:Accession: T40856

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-572 <RAM>

A:Cross-references: EMBL:AL21783; PIDN:CA857847.1; GSPDB:GN00069; SPDB:SPCC11E10.05C

A:Experimental source: strain 972h; cosmid c11E10

C:Genetics:

A:Gene: SPDB:SPCC11E10.05C

A:Map position: 3

Query Match 15.4%; Score 347; DB 2; Length 572;

Best Local Similarity 25.5%; Pred. No. 6.5e-21;

Matches 115; Conservative 67; Mismatches 157; Indels 112; Gaps 16;

Qy 49 YGIMFDAGSTGTRIHYVTF-----VQKMPQQLPILGEVFD-----SVKPGLSAFVD 95

Db 5 YGIFDAGSSSRLLIYNDYDTSLSKVKKLELIETGIGDGGKSLKVPGLSSPAN 64

Qy 96 QPKQ-GAETVOGLEVAKDSIPRSHWKTTPVVKATAGRLLL-PEHAKALLFEVKEIFR 153

Db 65 NPKHVKKHLKELLFAAHPKOVHKTPTVFLSATAGVLLGVDAQNKILSHACRYIK 124

Qy 154 KSPFLVP--KGSVSTMDGSDGILAWTVNFLTGLGHGRQETVCTLDLGASTQITFLP 211

Db 125 NYDFDPNCSNIRVIDGAEGNYGLATNVLKTLSEKDTSTVGFLDXGASVQLAF-- 182

Qy 212 QPEKTEQTPRGYLTSEFVNST-----YKLYTHSYLGFGLKAARLATLAL- 258

Db 183 -----ELPPSOLKXKDSISTVHIGLQNGOOLEYPLFVITWTLGFGANEAYRVLGLLI 235

Qy 259 ETE-CTDGTGHTFRSACLPMLEAEWIFGKVKYQYQ-----NQSSEVGFEPY 304

Db 236 ESENKVGNTLSDPCSLR--GRTYDIDGIEFAGTGDLKQCLKLTNLLNKDKPCSMDCPN 293

Qy 305 AEVLVVRVGRKLQPEVEVQSGYAFSYYYDRAVDMDIDYEKGGILKVEDFERKAREVC- 363

Db 294 PDGISI-----PPVDFANTEPVGVSEFYNTNDV-----FDMGGSYHFPNFKKVDYCG 343

Qy 364 -----DNLENFTSGSPFLCMDSYITALLKQDGFADSTV----- 398

Db 344 TEWETMLSRNLNKELTPSTDENKLEK-----LCFKASWALNVLHGFQVFKNTSSND 396

Qy 399 -----LQTKKKNNIETGWAIG 415

Db 397 AKDGLSVIPAYHSPFTSLKXIERTEVSWTLG 427

RESULT 13

E86276

hypothetical protein F4LI17.1 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001

C:Accession: E86276

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alcocin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, ansen, N.P.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzi Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tall ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: E86276

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-405 <STO>

A:Cross-references: GB:AE005172; NID:G7262666; PIDN:AAF43924.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

Query Match 14.6%; Score 329.5; DB 2; Length 405;

Best Local Similarity 26.3%; Pred. No. 1.2e-19;

Matches 110; Conservative 63; Mismatches 158; Indels 87; Gaps 15;

Qy 48 LYGIMFDAGSTGTRIHYVTFVQKMPQQLPILGEVFDVSKVGLSAFVDQPKGAETVOGL 107

Db 3 VFGYWFESGKP-----VFDFGEHYASL-----KLSFGLSSVADNPEGASVYTKL 48

Qy 108 LEVAKDSIPRSHWKTTPVVKATAGRLLL-PEHAKALLFEVKEIFRSPPLVPGKSVSIM 167

Db 49 VEFAGRIKPKGLKSDIRLMTATAGRLLDVPPVQEQILDVTRVLRSSGKFKQDEWATVI 108

Qy 168 DGSDEGLLAWTVNFLTGLGHGRQETVCTLDLGASTQITLPOPEKTELETPRGYLT 227

Db 109 SGTDDEGLIYAWVANHAGLSGGDPLKTTGIVELGGASQVTFVS-----EHVPPPEFSRT 163

Qy 228 FEMFNSTYKLYTHSYLGFGLKAARLATLGALETE---GTDGHTFRSACLFR----- 275

Db 164 ISYGNVSYTYIYSHSFLDQDAEDKLLESLSQNSVAASTGDGIVEDPCTKGYIYDTHSQ 223

Qy 276 -----MLEAEWIF--GGVKYQYQNGEVEGFECYAEVLVVR--GKLEQPEEVQSGSYFA 328

Db 224 KDSGGLSESKFKASLQVQAAGD-----FTKRSATLAWLQEGKEN----- 265

Qy 329 FSVYVDRAVDTMDIVEKGGILKVEDF-----ERRKREVCNKL 366

Db 266 -CAYKCSIGSTTPNIQGSFLATENFHTSKFGLGERKEMLESEMILAGKRFQGEWSKL 324

Qy 367 -ENP-TSGSPFL---CMDSYITALLKQDGFPA-DSTVLQLTKKV--NNIETGWAIGA 416

Db 325 KEKYPITKDKYLHRYCFSSAYIIISMLHSDLSGVALDDERIKYASKAGKENIPLDWAIGA 382

RESULT 14

T04439

hypothetical protein T18B16.150 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 30-Apr-1999

C:Accession: T04439

R:Baran, M.; Benes, V.; Rechmann, S.; Borkova, D.; Ansgorge, W.; Bancroft, I.; Mewes, submitted to the Protein Sequence Database, April 1998

A:Reference number: Z15359

A;Accession: T04439
A;Molecule type: DNA
A;Residues: 1-1052 <BEV>
A;Cross-references: EMBL:AL021687
A;Experimental source: cultivar Columbia; BAC clone T18B16
C;Genetics:
A;Map position: 4
A;Intons: 79/3; 118/3; 217/1; 295/3; 396/3; 531/2; 815/3
A;Note: T18B16.150

Query Match 12.2%; Score 274; DB 2; Length 1052;
Best Local Similarity 24.3%; Pred. No. 2.1e-14;
Matches 107; Conservative 70; Mismatches 171; Indels 82; Gaps 19;
QY 51 IMFDASTGTGRIHVY--TFVQKMPGQLPIL-----EGEVFD--SVKPGLSAFV 94
DB 525 LVIVSITGTRAVYQASINVKDSSLPIVMKSLTEGSKSRGRAYDRMETEPGDKLV 584
QY 95 DQPKQAGET-VQGLLEVAKDSIPRSHWKTTPVVKATAGRLILPEKAKALLPEVKEIFR 153
DB 585 NN-RTGLKTAIKELIQWASKQIPKNAHRTTSLFVYATAGVRLRPADSSMILGNVMSILA 643
QY 154 KSPFLVPKGSVMDSDEGILAWVTNFLTGLHG-HROETVGTLDLGGASTQITFLPQ 212
DB 644 KSPFTCRREWVKIISGTEEAFFGWTALNYQTSMLGALPKKATFGALDLGGSSLVQTFENE 703
QY 213 PEKTLSTPRGYTSTFEMFNSTYKLYTHSVLGFGLKAA-----RLATL----- 255
DB 704 -ERTHNETN---LNLKIGSVNHLAYSAGYGLNDADFRSVVHLLKXLPNVNKSDLIE 758
QY 256 GALETEGTGHTFRSACLPRWLEAWTFGGVKYQYGNQCEGVGFE-----PCYAEVLR 309
DB 759 GKLE-----MKHPLNSGNGOYICSCASSVQGGKKGSGVSIKLVGAENWGECSA 810
QY 310 VVR-GKXHOPEEVR--GSYAFSYVYDRAVDTMDIDYKGGI---LKVEDFERKAREVC 363
DB 811 LAXNAPCALPDGYPRHGQYAVSGFF-----VVYRPFNLSEASLDDVLEKGRFC 862
QY 364 DN-----LENFTSGSPEL---CMDSLXITALLKDGFGPADSTV-----LQLTKKVNNET 410
DB 863 DKAWQARTSVSPQPIEQYCFRAPHYVLSIREGLYITDKQIIGSGSIWTLGVALLS 922
QY 411 GWALGATFHL 420
DB 923 GKALSSTLGL 932

RESULT 15

C86276
7A19.33 protein - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 18-Nov-2002
C;Accession: C86276
R;neologs, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: C86276
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-508 <STO>
A;Cross-references: GB:AE005172; NID:95080800; PIDN:AAD39310.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1
C;Superfamily: nucleoside triphosphatase chromatin-associated

Query Match 11.0%; Score 248; DB 2; Length 508;
Best Local Similarity 23.3%; Pred. No. 1.1e-12;
Matches 110; Conservative 77; Mismatches 159; Indels 126; Gaps 21;
QY 12 LVVSCVCSAV-----SHRQQTWFEIGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
DB 35 VIVACVTIALGLLFTGYILRSGRNR-----VSLHYSVIIDGSSGT 77
QY 61 RIHVYTFVQKMPGQLPIL--GEVFD--SVKPGLSAFV--PGLSAFVDQPKQAGTVOGLLEVAKDSIP 116
DB 78 RVHVEGY--RIESGKPVDFGENTYASLKLSPGLSAYADNPEGVSESVTELVEFAKRVH 135
QY 117 RSHWKTTPVVKATAGRLILPEKAKALLPEVKEIFRSPPLVPPKGSVIMDSGB--GI 174
DB 136 KGLKXSDIRLWATAGMLLELPVQEQILDVTRVLRSGF-----DFRDEWASV 185
QY 175 LAWVTNFLTGLHGHROETVGTLDLG-----GASTQITFLPQPEKTLSTPRGYLTPSP 228
DB 186 ISEILENF-----QDLMKVYMLGLLLIMRSVRLEVTFF-----STELVPSBFSRTL 231
QY 229 EMFNSTYKLYTHSVLGFGLKAA RLATL-----ETEGTGHTRFSA CLPRWLEAWI--- 282
DB 232 AYGNVSYNLYSHSFLDPQDAQAEKLSLXNSAANSTGEGIVDPFCIPKGYILETNLQK 291
QY 283 ----FGVKYQYGNQCEGVGFEPCYAEVLRVVR-----GKXHOPEEVRQGS 325
DB 292 DLPGFLADKGTATLQAGNFSCKSAFAMLOBEKKGCTYKRCSIGSIPTFN--LQGS 349
QY 326 FYAFSYVY-----DRAVDTMDI-----DYKGGILKVEDFERKAREVCNDLE 367
DB 350 FLATENFHTSKXFFGLGEKEMLSMILAGKRCFGGEWSK---LKVYPTFFD-----ENLL 402
QY 368 NTFSGSPFLCMLSXITALLKDGFGFA-DSTVLQUTKKV--NNIETGVALGA 416
DB 403 RY-----CPSAIIISMLHSLGVALDDERIKYASKAGEEDIPLDWALGA 447

Search completed: July 1, 2004, 13:49:25
Job time : 18.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2004, 13:48:47 ; Search time 44.5 Seconds

(without alignments)
2993.921 Million cell updates/sec

Title: US-10-091-085-7

Perfect score: 2252

Sequence: 1 MATSWGTFFMLVVCVCSA.....ETGWCALGATFLLQLSLGISH 428

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1276540 seqs, 311283616 residues

Total number of hits satisfying chosen parameters: 1276540

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA:*
- 1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/prodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/prodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
 - 10: /cgn2_6/prodata/1/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/prodata/1/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/prodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 17: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 18: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2252	100.0	428	13	US-10-091-085-7
2	2252	100.0	428	13	US-10-092-063-7
3	2252	100.0	428	14	US-10-286-926-7
4	2235	99.2	428	13	US-10-091-085-3
5	2235	99.2	428	13	US-10-091-085-5
6	2235	99.2	428	13	US-10-092-063-3
7	2235	99.2	428	13	US-10-092-063-5
8	2235	99.2	428	14	US-10-286-926-3
9	2235	99.2	428	14	US-10-286-926-5
10	2235	99.2	428	16	US-10-231-913-127
11	2221	98.6	428	15	US-10-408-765A-2296
12	2089	92.8	405	13	US-10-092-063-25
13	2089	92.8	405	14	US-10-286-926-25
14	1979.5	87.9	427	15	US-10-231-913-126
15	1822.5	80.9	465	13	US-10-092-063-39

Sequence 876, App
Sequence 876, App
Sequence 27, Appl
Sequence 27, Appl
Sequence 123, Appl
Sequence 36, Appl
Sequence 271, Appl
Sequence 124, Appl
Sequence 38, Appl
Sequence 125, App
Sequence 6417, App
Sequence 4, Appl
Sequence 3, Appl
Sequence 2169, App
Sequence 19, Appl
Sequence 45875, A
Sequence 9, Appl
Sequence 272, App
Sequence 132351, A
Sequence 230158, A
Sequence 1713, App
Sequence 15, Appl
Sequence 36762, A
Sequence 49933, A
Sequence 151710, A
Sequence 51762, A
Sequence 145076, A
Sequence 204232, A
Sequence 37241, A

330 9 US-09-925-299-876
330 10 US-09-925-299-876
484 13 US-10-092-063-27
484 14 US-10-286-926-27
484 15 US-10-231-913-123
467 15 US-10-231-913-36
379 15 US-10-231-913-271
484 15 US-10-231-913-124
446 15 US-10-231-913-38
455 15 US-10-231-913-125
479 15 US-10-369-493-6447
476 9 US-09-835-147-4
476 9 US-09-835-147-3
556 15 US-10-369-493-2169
467 9 US-09-129-112-19
467 12 US-10-425-114-45875
459 9 US-09-129-112-9
410 15 US-10-231-913-272
534 16 US-10-437-963-132351
467 12 US-10-424-599-230158
462 9 US-09-129-112-2
518 15 US-10-369-493-1713
462 9 US-09-129-112-15
433 12 US-10-425-114-36762
443 12 US-10-425-114-49933
472 16 US-10-437-963-151710
457 12 US-10-425-114-51762
459 12 US-10-424-599-145076
449 16 US-10-437-963-204232
465 12 US-10-425-114-37241

ALIGNMENTS

RESULT 1

US-10-091-085-7
; Sequence 7, Application US/10091085
; Publication NO. US20020146772A1
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 28110/35761
; CURRENT APPLICATION NUMBER: US/10/091,085
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-091-085-7

Query Match 100.0%; Score 2252; DB 13; Length 428;
Best Local Similarity 100.0%; Pred. No. 3.7e-218;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MATSWGTFFMLVVCVCSAVSHRQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
1 MATSWGTFFMLVVCVCSAVSHRQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

QY 61 RIHVYTFVQKMPGQQLPILGEVFDVSKPGLSAFVDQPKQAGTVOGLLEVAKDSIPRSHW 120
DB 61 RIHVYTFVQKMPGQQLPILGEVFDVSKPGLSAFVDQPKQAGTVOGLLEVAKDSIPRSHW 120
QY 121 KKTPEVLKATAGLRLLPEHKAALLFEVKIEPRKSPFLVPKGSVSIITQDEGFIWVTV 180
DB 121 KKTPEVLKATAGLRLLPEHKAALLFEVKIEPRKSPFLVPKGSVSIITQDEGFIWVTV 180
QY 181 NFLTQGLHGHROETVGTDLGGASTQITLPOFEKTLBOTPRGYLTSFEMFNSTYKLYTH 240
DB 181 NFLTQGLHGHROETVGTDLGGASTQITLPOFEKTLBOTPRGYLTSFEMFNSTYKLYTH 240
QY 241 SYLGFGLKAARLATIAGLETGDTGHTFRSACLPRWLEAEWIFGCVKYQYGGNQEVEVP 300
DB 241 SYLGFGLKAARLATIAGLETGDTGHTFRSACLPRWLEAEWIFGCVKYQYGGNQEVEVP 300
QY 301 EPCYAEVLVRVRKGLHQPEVQSGSFYAFSYVYDRAVDTMDIDYKGGILKVEDFERKAR 360
DB 301 EPCYAEVLVRVRKGLHQPEVQSGSFYAFSYVYDRAVDTMDIDYKGGILKVEDFERKAR 360
QY 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQJTKKVNNIETGALGATFHL 420
DB 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQJTKKVNNIETGALGATFHL 420
QY 421 LQSLGISH 428
DB 421 LQSLGISH 428

RESULT 2

US-10-092-063-7
; Sequence 7, Application US/10092063
; Publication No. US20020173005A1
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES
; FILE REFERENCE: 28110/35908
; CURRENT APPLICATION NUMBER: US/10/092,063
; CURRENT FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-063-7

Query Match 100.0%; Score 2252; DB 13; Length 428;
Best Local Similarity 100.0%; Pred. No. 3.7e-218;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATSWGTVPFMLVVCVCSAVSHRNQQTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
DB 1 MATSWGTVPFMLVVCVCSAVSHRNQQTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
QY 61 RIHVYTFVQKMPGQQLPILGEVFDVSKPGLSAFVDQPKQAGTVOGLLEVAKDSIPRSHW 120
DB 61 RIHVYTFVQKMPGQQLPILGEVFDVSKPGLSAFVDQPKQAGTVOGLLEVAKDSIPRSHW 120

QY 121 KKTPEVLKATAGLRLLPEHKAALLFEVKIEPRKSPFLVPKGSVSIITQDEGFIWVTV 180
DB 121 KKTPEVLKATAGLRLLPEHKAALLFEVKIEPRKSPFLVPKGSVSIITQDEGFIWVTV 180
QY 181 NFLTQGLHGHROETVGTDLGGASTQITLPOFEKTLBOTPRGYLTSFEMFNSTYKLYTH 240
DB 181 NFLTQGLHGHROETVGTDLGGASTQITLPOFEKTLBOTPRGYLTSFEMFNSTYKLYTH 240
QY 241 SYLGFGLKAARLATIAGLETGDTGHTFRSACLPRWLEAEWIFGCVKYQYGGNQEVEVP 300
DB 241 SYLGFGLKAARLATIAGLETGDTGHTFRSACLPRWLEAEWIFGCVKYQYGGNQEVEVP 300
QY 301 EPCYAEVLVRVRKGLHQPEVQSGSFYAFSYVYDRAVDTMDIDYKGGILKVEDFERKAR 360
DB 301 EPCYAEVLVRVRKGLHQPEVQSGSFYAFSYVYDRAVDTMDIDYKGGILKVEDFERKAR 360
QY 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQJTKKVNNIETGALGATFHL 420
DB 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQJTKKVNNIETGALGATFHL 420
QY 421 LQSLGISH 428
DB 421 LQSLGISH 428

RESULT 3

US-10-286-926-7
; Sequence 7, Application US/10286926
; Publication No. US20030175752A1
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
; FILE REFERENCE: 28110/36457CON
; CURRENT APPLICATION NUMBER: US/10/286,926
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/122449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-286-926-7

Query Match 100.0%; Score 2252; DB 14; Length 428;
Best Local Similarity 100.0%; Pred. No. 3.7e-218;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATSWGTVPFMLVVCVCSAVSHRNQQTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
DB 1 MATSWGTVPFMLVVCVCSAVSHRNQQTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

[illegible]

RESULT 4
US-10-091-085-3
; Sequence 3, Application US/10091085
; Publication No. US20020146772A1
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE

Query Match	99.2%	Score 2235;	DB 13;	Length 428;
Best Local Similarity	99.3%	Pred. No. 1.9e-216;		
Matches 425;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
OY	1	MATSGTFFPMLVVCVCSAVSHENQQTWTEGIFLSSMCPINVSASTLYGINFMDAGSTGT	60	
Db	1	MATSGTFFPMLVVCVCSAVSHENQQTWTEGIFLSSMCPINVSASTLYGINFMDAGSTGT	60	
OY	61	RIHVYTFVQKMPGQLPILEGEVFDVSXKGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW	120	
Db	61	RIHVYTFVQKMPGQLPILEGEVFDVSXKGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW	120	
OY	121	KKTPVVLKATAGLRLLPEHKAKALLFEVKGIIFKKSPLFLPKRGSVSIIMTQDDEGIFAMVTV	180	

Db	121	KKT	PVVLKATAGLRLLPEHKA	KALLFEVKEIFRKSPPFLVRKGSVSI	MDGSDGII	LAWTV	180
Qy	181	NFLT	QOLGHRHQETVGTGLDLGGASTQITFLPQ	PEKTLSEQTPRGYLTSPEMFNSTYKLYTH	240		
Db	181	NFLT	QOLGHRHQETVGTGLDLGGASTQITFLPQ	PEKTLSEQTPRGYLTSPEMFNSTYKLYTH	240		
Qy	241	SYL	FGGLKAARLATLGALETGTGDTGHTFRSACL	PRWLEAWIFGKVKYOGGNORGEVGF	300		
Db	241	SYL	FGGLKAARLATLGALETGTGDTGHTFRSACL	PRWLEAWIFGKVKYOGGNORGEVGF	300		
Qy	301	EPC	YAEVLVRVVRGKLHQPEEVQORGSFYAFS	YYYDRAVDTMDIYEKGKILKVEDFERKAR	360		
Db	301	EPC	YAEVLVRVVRGKLHQPEEVQORGSFYAFS	YYYDRAVDTMDIYEKGKILKVEDFERKAR	360		
Qy	361	EVC	DNLENFTSGSPFLCMDLSYITALLKDG	GFADSTVLQITKKVNNIETGWCATPHL	420		
Db	361	EVC	DNLENFTSGSPFLCMDLSYITALLKDG	GFADSTVLQITKKVNNIETGWCATPHL	420		
Qy	421	LOS	LIGISH	428			
Db	421	LOS	LIGISH	428			

RESULT 5
US-10-091-085-5
Sequence 5, Application US/10091085
Publication No. US20020146772A1
GENERAL INFORMATION:
APPLICANT: Ford, John
APPLICANT: Mulero, Julio
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
POLYPEPTIDES
TITLE OF INVENTION: POLYPEPTIDES
FILE REFERENCE: 28110/35761
CURRENT APPLICATION NUMBER: US/10/091,085
CURRENT FILING DATE: 2002-03-05
PRIOR APPLICATION NUMBER: 09/350,836
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 09/273,447
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 09/118,205
PRIOR FILING DATE: 1998-07-15
PRIOR APPLICATION NUMBER: 09/122,449
PRIOR FILING DATE: 1998-07-24
PRIOR APPLICATION NUMBER: 09/244,444
PRIOR FILING DATE: 1999-02-04
NUMBER OF SEQ ID NOS: 23
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 5
LENGTH: 428
TYPE: PRT
ORGANISM: Homo sapiens
US-10-091-085-5

Query Match	99.3%	Score 2235;	DB 13;	Length 428;
Best Local Similarity	99.3%	Pred. No. 1.9e-216;		
Matches 425;	Conservative	0;	Mismatches 3;	Indels 0;
Gaps	0			

QY 241 SYLGFLKAARLATIAGLETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGVP 300
DB 241 SYLGFLKAARLATIAGLETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGVP 300
QY 301 EPCYAEVLVRVVKGLHQPBEVQSGFYAFSYYYDRAVDMDIYEKGGILKVEDFERKAR 360
DB 301 EPCYAEVLVRVVKGLHQPBEVQSGFYAFSYYYDRAVDMDIYEKGGILKVEDFERKAR 360
QY 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQTKKVNNIETGALGATFHL 420
DB 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQTKKVNNIETGALGATFHL 420
QY 421 LQSLGISH 428
DB 421 LQSLGISH 428

RESULT 6

US-10-092-063-3
; Sequence 3, Application US/10092063
; Publication No. US20020173005A1
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES
; FILE REFERENCE: 28110/35908
; CURRENT APPLICATION NUMBER: US/10/092,063
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; TYPE: PRT
; LENGTH: 428
; ORGANISM: Homo sapiens
US-10-092-063-3

Query Match 99.2%; Score 2235; DB 13; Length 428;
Best Local Similarity 99.3%; Pred. No. 1.9e-216;
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MATSWGTVFFMLVVCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
DB 1 MATSWGTVFFMLVVCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
QY 61 RIHVYTFVQKPGQLPILGEVFDVSKPGLSAFVDQPKGAETVQGLLEVAKDSIPRSHW 120
DB 61 RIHVYTFVQKPGQLPILGEVFDVSKPGLSAFVDQPKGAETVQGLLEVAKDSIPRSHW 120
QY 121 KKTVPVLKATAGRLLPHEKAKALLFEVKEIPRKSPLVPKGSVIMTGDEGIFAWTV 180
DB 121 KKTVPVLKATAGRLLPHEKAKALLFEVKEIPRKSPLVPKGSVIMTGDEGIFAWTV 180
QY 181 NFLTQQLHGRQETVGTLDLGGASTQITFLPQFKEKLEQTPRGYLTSEFMENSTYKLYTH 240
DB 181 NFLTQQLHGRQETVGTLDLGGASTQITFLPQFKEKLEQTPRGYLTSEFMENSTYKLYTH 240
QY 241 SYLGFLKAARLATIAGLETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGVP 300
DB 241 SYLGFLKAARLATIAGLETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGVP 300

QY 301 EPCYAEVLVRVVKGLHQPBEVQSGFYAFSYYYDRAVDMDIYEKGGILKVEDFERKAR 360
DB 301 EPCYAEVLVRVVKGLHQPBEVQSGFYAFSYYYDRAVDMDIYEKGGILKVEDFERKAR 360
QY 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQTKKVNNIETGALGATFHL 420
DB 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQTKKVNNIETGALGATFHL 420
QY 421 LQSLGISH 428
DB 421 LQSLGISH 428

RESULT 7

US-10-092-063-5
; Sequence 5, Application US/10092063
; Publication No. US20020173005A1
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES
; FILE REFERENCE: 28110/35908
; CURRENT APPLICATION NUMBER: US/10/092,063
; CURRENT FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; TYPE: PRT
; LENGTH: 428
; ORGANISM: Homo sapiens
US-10-092-063-5

Query Match 99.2%; Score 2235; DB 13; Length 428;
Best Local Similarity 99.3%; Pred. No. 1.9e-216;
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MATSWGTVFFMLVVCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
DB 1 MATSWGTVFFMLVVCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
QY 61 RIHVYTFVQKPGQLPILGEVFDVSKPGLSAFVDQPKGAETVQGLLEVAKDSIPRSHW 120
DB 61 RIHVYTFVQKPGQLPILGEVFDVSKPGLSAFVDQPKGAETVQGLLEVAKDSIPRSHW 120
QY 121 KKTVPVLKATAGRLLPHEKAKALLFEVKEIPRKSPLVPKGSVIMTGDEGIFAWTV 180
DB 121 KKTVPVLKATAGRLLPHEKAKALLFEVKEIPRKSPLVPKGSVIMTGDEGIFAWTV 180
QY 181 NFLTQQLHGRQETVGTLDLGGASTQITFLPQFKEKLEQTPRGYLTSEFMENSTYKLYTH 240
DB 181 NFLTQQLHGRQETVGTLDLGGASTQITFLPQFKEKLEQTPRGYLTSEFMENSTYKLYTH 240
QY 241 SYLGFLKAARLATIAGLETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGVP 300
DB 241 SYLGFLKAARLATIAGLETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGVP 300
QY 301 EPCYAEVLVRVVKGLHQPBEVQSGFYAFSYYYDRAVDMDIYEKGGILKVEDFERKAR 360

Db 301 EPCYAEVLVRGKLGHPVEVQSGFYAFSYYDDRAVDTMDIDYKGGILKVEDPERKAR 360
Qy 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTKKVNNIETGHALGATPHL 420
Db 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTKKVNNIETGHALGATPHL 420
Qy 421 LOSLGISH 428
Db 421 LOSLGISH 428

RESULT 8
US-10-286-926-3
; Sequence 3, Application US/10286926
; Publication No. US20030175752A1
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
; FILE REFERENCE: 28110/36457CON
; CURRENT APPLICATION NUMBER: US/10/286,926
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/122449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-286-926-3

Query Match 99.2%; Score 2235; DB 14; Length 428;
Best Local Similarity 99.3%; Pred. No. 1.9e-216;
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 MATSWGTVFPMVLVSVCSAVSHRNQOTWFEGLFSLSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSWGTVFPMVLVSVCSAVSHRNQOTWFEGLFSLSMCPINVSASTLYGIMFDAGSTGT 60
Qy 61 RIHYVTVFQKMPGQLPILGEVDSVPKGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
Db 61 RIHYVTVFQKMPGQLPILGEVDSVPKGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
Qy 121 KKTPEVVKATAGLRLPEHAKALLFEVKIIFRKSPPFLVPKGSVIMTGDGFIWVTV 180
Db 121 KKTPEVVKATAGLRLPEHAKALLFEVKIIFRKSPPFLVPKGSVIMTGDGFIWVTV 180
Qy 181 NFLTGQLHGRHRETQVGTLDLGGASTQITFLPQFEKTLQTPRGYLTSPFEMNSTYKLYTH 240
Db 181 NFLTGQLHGRHRETQVGTLDLGGASTQITFLPQFEKTLQTPRGYLTSPFEMNSTYKLYTH 240
Qy 241 SYLGFGKAAFLATLGALETGDTGHTFRSACLPRLWEAEWIFGVKYYQYGGNQEVEGV 300
Db 241 SYLGFGKAAFLATLGALETGDTGHTFRSACLPRLWEAEWIFGVKYYQYGGNQEVEGV 300

Qy 301 EPCYAEVLVRGKLGHPVEVQSGFYAFSYYDDRAVDTMDIDYKGGILKVEDPERKAR 360
Db 301 EPCYAEVLVRGKLGHPVEVQSGFYAFSYYDDRAVDTMDIDYKGGILKVEDPERKAR 360
Qy 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTKKVNNIETGHALGATPHL 420
Db 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTKKVNNIETGHALGATPHL 420
Qy 421 LOSLGISH 428
Db 421 LOSLGISH 428

RESULT 9
US-10-286-926-5
; Sequence 5, Application US/10286926
; Publication No. US20030175752A1
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
; FILE REFERENCE: 28110/36457CON
; CURRENT APPLICATION NUMBER: US/10/286,926
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/122449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 5
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-286-926-5

Query Match 99.2%; Score 2235; DB 14; Length 428;
Best Local Similarity 99.3%; Pred. No. 1.9e-216;
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 MATSWGTVFPMVLVSVCSAVSHRNQOTWFEGLFSLSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSWGTVFPMVLVSVCSAVSHRNQOTWFEGLFSLSMCPINVSASTLYGIMFDAGSTGT 60
Qy 61 RIHYVTVFQKMPGQLPILGEVDSVPKGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
Db 61 RIHYVTVFQKMPGQLPILGEVDSVPKGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
Qy 121 KKTPEVVKATAGLRLPEHAKALLFEVKIIFRKSPPFLVPKGSVIMTGDGFIWVTV 180
Db 121 KKTPEVVKATAGLRLPEHAKALLFEVKIIFRKSPPFLVPKGSVIMTGDGFIWVTV 180
Qy 181 NFLTGQLHGRHRETQVGTLDLGGASTQITFLPQFEKTLQTPRGYLTSPFEMNSTYKLYTH 240
Db 181 NFLTGQLHGRHRETQVGTLDLGGASTQITFLPQFEKTLQTPRGYLTSPFEMNSTYKLYTH 240

Qy 241 SYLGFLGKAARLATLGALETGTGHTFRSACLPRLWEAEWIFGGVVKYQYGGNQEVEVGF 300
 Db 241 SYLGFLGKAARLATLGALETGTGHTFRSACLPRLWEAEWIFGGVVKYQYGGNQEVEVGF 300
 Qy 301 EPCVAEVLVRVGRKLGHPSEVQVGSFYAFSYDDRAVDTDMDYKGGILKVEDFERKAR 360
 Db 301 EPCVAEVLVRVGRKLGHPSEVQVGSFYAFSYDDRAVDTDMDYKGGILKVEDFERKAR 360
 Qy 361 EVCDNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQTLTKKNNIETGWMALGATFHL 420
 Db 361 EVCDNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQTLTKKNNIETGWMALGATFHL 420
 Qy 421 LQSLGISH 428
 Db 421 LQSLGISH 428

RESULT 10

US-10-231-913-127
 ; Sequence 127, Application US/10231913
 ; Publication No. US20040005576A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Guo, Xiaojia S.
 ; APPLICANT: Li, Li
 ; APPLICANT: Patturajan, Meera
 ; APPLICANT: Shukets, Richard A.
 ; APPLICANT: Casman, Stacie J.
 ; APPLICANT: Malyanekar, Uriel M.
 ; APPLICANT: Tchernev, Velizar T.
 ; APPLICANT: Vernet, Corine A.
 ; APPLICANT: Spytek, Kimberly A.
 ; APPLICANT: Shenoy, Suresh G.
 ; APPLICANT: Alsobrook II, John P.
 ; APPLICANT: Edinger, Schlomit
 ; APPLICANT: Peyman, John A.
 ; APPLICANT: Stone, David J.
 ; APPLICANT: Ellerman, Karen
 ; APPLICANT: Gangolli, Esha A.
 ; APPLICANT: Boldog, Ference L.
 ; APPLICANT: Colman, Steven D.
 ; APPLICANT: Eisen, Andrew J.
 ; APPLICANT: Liu, Xiaohong
 ; APPLICANT: Padigaru, Muralidhara
 ; APPLICANT: Spaderna, Steven K.
 ; APPLICANT: Zerhusen, Bryan D.
 ; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 21402-216
 ; CURRENT APPLICATION NUMBER: US/10/231,913
 ; CURRENT FILING DATE: 2002-08-30
 ; PRIOR APPLICATION NUMBER: 60/251,660
 ; PRIOR FILING DATE: 2000-12-06
 ; PRIOR APPLICATION NUMBER: 60/255,029
 ; PRIOR FILING DATE: 2000-12-12
 ; PRIOR APPLICATION NUMBER: 60/260,326
 ; PRIOR FILING DATE: 2001-01-08
 ; PRIOR APPLICATION NUMBER: 60/263,800
 ; PRIOR FILING DATE: 2001-01-24
 ; PRIOR APPLICATION NUMBER: 60/269,942
 ; PRIOR FILING DATE: 2001-02-20
 ; PRIOR APPLICATION NUMBER: 60/286,183
 ; PRIOR FILING DATE: 2001-04-24
 ; PRIOR APPLICATION NUMBER: 60/313,627
 ; PRIOR FILING DATE: 2001-08-20
 ; PRIOR APPLICATION NUMBER: 60/318,712
 ; PRIOR FILING DATE: 2001-09-12
 ; NUMBER OF SEQ ID NOS: 292
 ; SOFTWARE: Patent in Ver. 2.1
 ; SEQ ID NO 127
 ; LENGTH: 428
 ; TYPE: PRF
 ; ORGANISM: Homo sapiens
 US-10-231-913-127

Query Match 99.2%; Score 2235; DB 15; Length 428;
 Best Local Similarity 99.3%; Pred. No. 1.9e-215; Indels 0; Gaps 0;
 Matches 425; Conservative 0; Mismatches 3;
 Qy 1 MATSWGTVPFVLMVVCVSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
 Db 1 MATSWGTVPFVLMVVCVSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
 Qy 61 RIHVYTFVQKMPGQLPILGEVDFSVKPGLSAFVDPKQGAETVQGLLEVAKDISPRSHW 120
 Db 61 RIHVYTFVQKMPGQLPILGEVDFSVKPGLSAFVDPKQGAETVQGLLEVAKDISPRSHW 120
 Qy 121 KKTVPVLKATAGLRLLPEHKAKALLFEVKSEIFRKSPFLVYKPGSVSIMDGSDEGILLAWTV 180
 Db 121 KKTVPVLKATAGLRLLPEHKAKALLFEVKSEIFRKSPFLVYKPGSVSIMDGSDEGILLAWTV 180
 Qy 181 NFLTGQLHGHROBTVTGTLGGASTQITFLPQFETKLEQTPRGYLTSPFEMFNSTYKLYTH 240
 Db 181 NFLTGQLHGHROBTVTGTLGGASTQITFLPQFETKLEQTPRGYLTSPFEMFNSTYKLYTH 240
 Qy 241 SYLGFLGKAARLATLGALETGTGHTFRSACLPRLWEAEWIFGGVVKYQYGGNQEVEVGF 300
 Db 241 SYLGFLGKAARLATLGALETGTGHTFRSACLPRLWEAEWIFGGVVKYQYGGNQEVEVGF 300
 Qy 301 EPCVAEVLVRVGRKLGHPSEVQVGSFYAFSYDDRAVDTDMDYKGGILKVEDFERKAR 360
 Db 301 EPCVAEVLVRVGRKLGHPSEVQVGSFYAFSYDDRAVDTDMDYKGGILKVEDFERKAR 360
 Qy 361 EVCDNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQTLTKKNNIETGWMALGATFHL 420
 Db 361 EVCDNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQTLTKKNNIETGWMALGATFHL 420
 Qy 421 LQSLGISH 428
 Db 421 LQSLGISH 428

RESULT 11

US-10-408-765A-2296
 ; Sequence 2296, Application US/10408765A
 ; Publication No. US20040101874A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ghosh, Soumitra S.
 ; APPLICANT: Fahy, Bojin D.
 ; APPLICANT: Zhang, Bing
 ; APPLICANT: Gibson, Bradford W.
 ; APPLICANT: Taylor, Steven W.
 ; APPLICANT: Glenn, Gary M.
 ; APPLICANT: Warnock, Dale B.
 ; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
 ; FILE REFERENCE: 660088.465
 ; CURRENT APPLICATION NUMBER: US/10/408,765A
 ; CURRENT FILING DATE: 2003-04-04
 ; NUMBER OF SEQ ID NOS: 3077
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2296
 ; LENGTH: 428
 ; TYPE: PRF
 ; ORGANISM: Homo sapiens
 US-10-408-765A-2296

Query Match 98.6%; Score 2221; DB 16; Length 428;
 Best Local Similarity 98.8%; Pred. No. 5e-215; Indels 0; Gaps 0;
 Matches 423; Conservative 0; Mismatches 5;
 Qy 1 MATSWGTVPFVLMVVCVSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
 Db 1 MATSWGTVPFVLMVVCVSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
 Qy 61 RIHVYTFVQKMPGQLPILGEVDFSVKPGLSAFVDPKQGAETVQGLLEVAKDISPRSHW 120
 Db 61 RIHVYTFVQKMPGQLPILGEVDFSVKPGLSAFVDPKQGAETVQGLLEVAKDISPRSHW 120

121 KKTPTVLKATAGLRLPEHAKALLFEVKEIFRKSPLVPKGSVSIIMDGSDGILAWTV 180
 121 NFLTQGLHGRQETVGTDLGGASTQITFLPOPEKTLQOTPRGYLTSPFENSTYKLYTH 240
 121 NFLTQGLHGRQETVGTDLGGASTQITFLPOPEKTLQOTPRGYLTSPFENSTYKLYTH 240
 181 NFLTQGLHGRQETVGTDLGGASTQITFLPOPEKTLQOTPRGYLTSPFENSTYKLYTH 240
 181 NFLTQGLHGRQETVGTDLGGASTQITFLPOPEKTLQOTPRGYLTSPFENSTYKLYTH 240
 241 SYLGFGGLKAARLALATLGALETGTGHTFRSACLPRWLEAEWIFGGVYQYGGNOGEVGF 300
 241 SYLGFGGLKAARLALATLGALETGTGHTFRSACLPRWLEAEWIFGGVYQYGGNOGEVGF 300
 301 EPCYAEVLVRVRGKQHQPVEVQSGSFYAFSYYYDRAVDTMIDYEKGGILKVEDFERKAR 360
 301 EPCYAEVLVRVRGKQHQPVEVQSGSFYAFSYYYDRAVDTMIDYEKGGILKVEDFERKAR 360
 361 EVCDNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQ 420
 361 EVCDNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQ 420
 421 LQSLGISH 428
 421 LQSLGISH 428

RESULT 12
 ; Sequence 25, Application US/10092063
 ; Publication No. US20020173005A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ford, John
 ; APPLICANT: Mulero, Julio
 ; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES
 ; FILE REFERENCE: 28110/35908
 ; CURRENT APPLICATION NUMBER: US/10/092,063
 ; CURRENT FILING DATE: 2002-03-05
 ; PRIOR FILING DATE: 2002-01-31
 ; PRIOR APPLICATION NUMBER: 09/370,265
 ; PRIOR FILING DATE: 2002-01-11
 ; PRIOR APPLICATION NUMBER: PCT/US99/16180
 ; PRIOR FILING DATE: 1999-07-15
 ; PRIOR APPLICATION NUMBER: 09/350,836
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR APPLICATION NUMBER: 09/273,447
 ; PRIOR FILING DATE: 1999-03-19
 ; PRIOR APPLICATION NUMBER: 09/244,444
 ; PRIOR FILING DATE: 1999-02-04
 ; PRIOR APPLICATION NUMBER: 09/122,449
 ; PRIOR FILING DATE: 1998-07-24
 ; PRIOR APPLICATION NUMBER: 09/118,205
 ; PRIOR FILING DATE: 1998-07-16
 ; NUMBER OF SEQ ID NOS: 39
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 25
 ; LENGTH: 405
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-092-063-25

Query Match 92.8%; Score 2089; DB 13; Length 405;
 Best Local Similarity 99.2%; Pred. No. 9.7e-202;
 Matches 397; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MATSGTGVFFMLVWSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
 DB 1 MATSGTGVFFMLVWSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
 QY 61 RIHYTVTFQKMPGQLPILGEVFDVSKPGLSAFVDPQKQGAETVQGLLEVAKDSIPRSHW 120
 DB 61 RIHYTVTFQKMPGQLPILGEVFDVSKPGLSAFVDPQKQGAETVQGLLEVAKDSIPRSHW 120
 QY 121 KKTPTVLKATAGLRLPEHAKALLFEVKEIFRKSPLVPKGSVSIIMDGSDGILAWTV 180

121 KKTPTVLKATAGLRLPEHAKALLFEVKEIFRKSPLVPKGSVSIIMDGSDGILAWTV 180
 181 NFLTQGLHGRQETVGTDLGGASTQITFLPOPEKTLQOTPRGYLTSPFENSTYKLYTH 240
 181 NFLTQGLHGRQETVGTDLGGASTQITFLPOPEKTLQOTPRGYLTSPFENSTYKLYTH 240
 241 SYLGFGGLKAARLALATLGALETGTGHTFRSACLPRWLEAEWIFGGVYQYGGNOGEVGF 300
 241 SYLGFGGLKAARLALATLGALETGTGHTFRSACLPRWLEAEWIFGGVYQYGGNOGEVGF 300
 301 EPCYAEVLVRVRGKQHQPVEVQSGSFYAFSYYYDRAVDTMIDYEKGGILKVEDFERKAR 360
 301 EPCYAEVLVRVRGKQHQPVEVQSGSFYAFSYYYDRAVDTMIDYEKGGILKVEDFERKAR 360
 361 EVCDNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQ 400
 361 EVCDNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQ 400

RESULT 13
 ; Sequence 25, Application US/10286926
 ; Publication No. US20030175752A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ford, John
 ; APPLICANT: Mulero, Julio
 ; APPLICANT: Yeung, George
 ; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
 ; FILE REFERENCE: 28110/36457CON
 ; CURRENT APPLICATION NUMBER: US/10/286,926
 ; CURRENT FILING DATE: 2002-11-01
 ; PRIOR FILING DATE: 2000-04-25
 ; PRIOR APPLICATION NUMBER: 09/557,800
 ; PRIOR FILING DATE: 2000-01-11
 ; PRIOR APPLICATION NUMBER: 09/481,238
 ; PRIOR FILING DATE: 2000-01-11
 ; PRIOR APPLICATION NUMBER: 09/370,265
 ; PRIOR FILING DATE: 1999-08-09
 ; PRIOR APPLICATION NUMBER: PCT/US99/16180
 ; PRIOR FILING DATE: 1999-07-16
 ; PRIOR APPLICATION NUMBER: 09/350836
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR APPLICATION NUMBER: 09/273447
 ; PRIOR FILING DATE: 1999-03-19
 ; PRIOR APPLICATION NUMBER: 09/122449
 ; PRIOR FILING DATE: 1998-07-24
 ; PRIOR APPLICATION NUMBER: 09/244444
 ; PRIOR FILING DATE: 1999-02-04
 ; PRIOR APPLICATION NUMBER: 09/118,205
 ; PRIOR FILING DATE: 1998-07-16
 ; NUMBER OF SEQ ID NOS: 54
 ; SOFTWARE: PatentIn ver. 2.0
 ; SEQ ID NO 25
 ; LENGTH: 405
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-286-926-25

Query Match 92.8%; Score 2089; DB 14; Length 405;
 Best Local Similarity 99.2%; Pred. No. 9.7e-202;
 Matches 397; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MATSGTGVFFMLVWSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
 DB 1 MATSGTGVFFMLVWSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
 QY 61 RIHYTVTFQKMPGQLPILGEVFDVSKPGLSAFVDPQKQGAETVQGLLEVAKDSIPRSHW 120
 DB 61 RIHYTVTFQKMPGQLPILGEVFDVSKPGLSAFVDPQKQGAETVQGLLEVAKDSIPRSHW 120
 QY 121 KKTPTVLKATAGLRLPEHAKALLFEVKEIFRKSPLVPKGSVSIIMDGSDGILAWTV 180
 DB 121 KKTPTVLKATAGLRLPEHAKALLFEVKEIFRKSPLVPKGSVSIIMDGSDGILAWTV 180

QY 181 NLTQLHGRHQRQETVGTDLGGASTQITFLPQPKTKLEQTPRGVLTSPFEMFNSTVKLYTH 240
 DB 181 NLTQLHGRHQRQETVGTDLGGASTQITFLPQPKTKLEQTPRGVLTSPFEMFNSTVKLYTH 240
 QY 241 SYLGGLKAAARLATLGALETGTGHTTFRSACLPRWLEAEWIFGVKYQYGGNORGEVGF 300
 DB 241 SYLGGLKAAARLATLGALETGTGHTTFRSACLPRWLEAEWIFGVKYQYGGNORGEVGF 300
 QY 301 EPCYAEVLVRVGRKHLQPEEVQVGSFYAFSYVYDRAVDMDIYKGGILKVEDFERKAR 360
 DB 301 EPCYAEVLVRVGRKHLQPEEVQVGSFYAFSYVYDRAVDMDIYKGGILKVEDFERKAR 360
 QY 361 EVCNLENFTSGSPFLCMLDSVITALLKDGFGFADSTVLQ 400
 DB 361 EVCNLENFTSGSPFLCMLDSVITALLKDGFGFADSTVLQ 400

RESULT 14

US-10-231-913-126
 ; Sequence 126, Application US/10231913
 ; Publication No. US20040005576A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Guo, Xiaojia S.
 ; APPLICANT: Li, Li
 ; APPLICANT: Patturajan, Meera
 ; APPLICANT: Shimkets, Richard A.
 ; APPLICANT: Casman, Stacie J.
 ; APPLICANT: Malyankar, Uriel M.
 ; APPLICANT: Tchernev, Velizar T.
 ; APPLICANT: Vernet, Corine A.
 ; APPLICANT: Spytek, Kimberly A.
 ; APPLICANT: Shenoy, Suresh G.
 ; APPLICANT: Alsobrook II, John P.
 ; APPLICANT: Edinger, Schlomit
 ; APPLICANT: Peyman, John A.
 ; APPLICANT: Stone, David J.
 ; APPLICANT: Ellerman, Karen
 ; APPLICANT: Gangolli, Esha A.
 ; APPLICANT: Boldog, Ference L.
 ; APPLICANT: Colman, Steven D.
 ; APPLICANT: Eisen, Andrew J.
 ; APPLICANT: Liu, Xiaohong
 ; APPLICANT: Padigaru, Muralidhara
 ; APPLICANT: Spaderna, Steven K.
 ; APPLICANT: Zehrusen, Bryan D.
 ; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 21402-216
 ; CURRENT APPLICATION NUMBER: US/10/231,913
 ; PRIOR FILING DATE: 2002-08-30
 ; PRIOR APPLICATION NUMBER: 60/251,660
 ; PRIOR FILING DATE: 2000-12-06
 ; PRIOR APPLICATION NUMBER: 60/255,029
 ; PRIOR FILING DATE: 2000-12-12
 ; PRIOR APPLICATION NUMBER: 60/260,326
 ; PRIOR FILING DATE: 2001-01-08
 ; PRIOR APPLICATION NUMBER: 60/263,800
 ; PRIOR FILING DATE: 2001-01-24
 ; PRIOR APPLICATION NUMBER: 60/269,942
 ; PRIOR FILING DATE: 2001-02-20
 ; PRIOR APPLICATION NUMBER: 60/286,183
 ; PRIOR FILING DATE: 2001-04-24
 ; PRIOR APPLICATION NUMBER: 60/313,627
 ; PRIOR FILING DATE: 2001-08-20
 ; PRIOR APPLICATION NUMBER: 60/318,712
 ; PRIOR FILING DATE: 2001-09-12
 ; NUMBER OF SEQ ID NOS: 292
 ; SOFTWARE: Patent in Ver. 2.1
 ; SEQ ID NO 126
 ; LENGTH: 427
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-10-231-913-126

Query Match 80.9%; Score 1822.5; DB 13; Length 465;
 Best Local Similarity 83.3%; Pred. No. 9.7e-175;

Query Match 87.9%; Score 1979.5; DB 15; Length 427;
 Best Local Similarity 87.6%; Pred. No. 1.2e-190;
 Matches 374; Conservative 24; Mismatches 28; Indels 1; Gaps 1;
 QY 1 MATSWGTVPFVPMVSVCSAVSHRNOQTWPEGIPLSSMCPINVSASTLYGIMFDAGSTGT 60
 DB 1 MATSWGAV-FMLIIACUGSTVTFYRQQTWPEGFVLSMCPINVSAGTFYGIMFDAGSTGT 59
 QY 61 RIHYTFVQVQMPQLPILEGEVFDVSKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120
 DB 60 RIHYTFVQVQMPQLPILEGEIFDVSVPKGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 119
 QY 121 KKTVPVVKATAGLRLPEHKAKALLEVEKEIFRKSPLVPKPSVSTMTGDEGIPKAVTV 180
 DB 120 ERTVPVVKATAGLRLPEKQKALLLEVEKEIFRNSPFLVPDGSVSTMDGSYEGILAWTV 179
 QY 181 NLTQLHGRHQRQETVGTDLGGASTQITFLPQPKTKLEQTPRGVLTSPFEMFNSTVKLYTH 240
 DB 180 NLTQLHGRHQRQETVGTDLGGASTQITFLPQPKTKLEQTPRGVLTSPFEMFNSTVKLYTH 239
 QY 241 SYLGGLKAAARLATLGALETGTGHTTFRSACLPRWLEAEWIFGVKYQYGGNORGEVGF 300
 DB 240 SYLGGLKAAARLATLGALEBAKGYDGTFRSACLPRWLEAEWIFGVKYQYGGNORGEVGF 299
 QY 301 EPCYAEVLVRVGRKHLQPEEVQVGSFYAFSYVYDRAVDMDIYKGGILKVEDFERKAR 360
 DB 300 EPCYAEVLVRVQVGLHQPPEVRSFYAFSYVYDRAADTHLIDYKGGVLYKVEDFERKAR 359
 QY 361 EVCNLENFTSGSPFLCMLDSVITALLKDGFGFADSTVLQTKKVNIEGTGALGATPHL 420
 DB 360 EVCNLFSSSSGSPFLCMLDVTITALLKDGFGFADGTLQLTKKVNIEGTGALGATPHL 419
 QY 421 LQSLGITS 427
 DB 420 LQSLGIT 426

RESULT 15

US-10-092-063-39
 ; Sequence 39, Application US/10092063
 ; Publication No. US20020173005A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ford, John
 ; APPLICANT: Mulero, Julio
 ; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDE
 ; FILE REFERENCE: 28110/35908
 ; CURRENT APPLICATION NUMBER: US/10/092,063
 ; CURRENT FILING DATE: 2002-03-05
 ; PRIOR FILING DATE: 2002-01-31
 ; PRIOR APPLICATION NUMBER: PCT/US99/16180
 ; PRIOR FILING DATE: 1999-07-16
 ; PRIOR APPLICATION NUMBER: 09/350,836
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR APPLICATION NUMBER: 09/273,447
 ; PRIOR FILING DATE: 1999-03-19
 ; PRIOR APPLICATION NUMBER: 09/244,444
 ; PRIOR FILING DATE: 1999-02-04
 ; PRIOR APPLICATION NUMBER: 09/122,449
 ; PRIOR FILING DATE: 1998-07-24
 ; PRIOR APPLICATION NUMBER: 09/118,205
 ; PRIOR FILING DATE: 1998-07-16
 ; NUMBER OF SEQ ID NOS: 39
 ; SOFTWARE: Patent in Ver. 2.0
 ; SEQ ID NO 39
 ; LENGTH: 465
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-10-092-063-39

Query Match 80.9%; Score 1822.5; DB 13; Length 465;
 Best Local Similarity 83.3%; Pred. No. 9.7e-175;

Matches		350;	Conservative	25;	Mismatches	42;	Indels	3;	Gaps	3;
Qy	1	MATSWGTVFPM	LVVSVCSVSHRNQOTW	EGIFLSSMCPINVSASTLYGIMFDAGSTGT	60					
Db	1	MATSWGAV-FMLIIACV	STVFYREQQTW	EGVFLSSMCPINVSAGTFYGIMFDAGSTGT	59					
Qy	61	RIHYTTFVQXMPOLP	ILEGEVDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW	120						
Db	60	RIHYTTFVQXTAGQLP	FEGEIFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW	119						
Qy	121	KTPVVLKATAGLRLL	PEHKAKALLPEVKEIFRKSPFLVPKGSVSMTGODEGIFAMVTV	180						
Db	120	ERTFVVLKATAGLRLL	PEKRAQALLLEVERIFKNSPFLVPDGSVSIMDGSYEGILAMVTV	179						
Qy	181	NFLTQQLHGRQSTV	GTGLDGGASTQITFLPQFEKTLQTPRGYLTSEFMFNSTYKLYTH	240						
Db	180	NFLTQQLHGRQSTV	GTGLDGGASTQITFLPQFEKTLQTPRGYLTSEFMFNSTFKLYTH	239						
Qy	241	SYLGFGKKAARLAT	GLGALETGDTGHTFRSACLPRWLEAENIFGGVKYQYGGNQEGEVGF	300						
Db	240	SYLGFGKKAARLAT	GLGALETGDTGHTFRSACLPRWLEAENIFGGVKYQYGGNQEGEMGF	299						
Qy	301	EPCYAEVLRVVRGKL	HQPEVRQGSFYAFSYYYDRAVDTMDIYERKGILKVEDFERKAR	360						
Db	300	EPCYAEVLRVVQVKL	HQPEBVRGSAFYAFSYYYDRAADTHLIDYERKGVILKVEDFERKAR	359						
Qy	361	EVCNLENFTSGSPFL	CMDSYITALLKDGFGADSTVLCITKKVNNIETGVALGATPHL	420						
Db	360	EVCNLSGFSGSGPFL	CMDLTYITALLKGLGFAERHPLTAHKE-SEQHRDW-LGLGZHL	417						

Search completed: July 1, 2004, 13:58:41
Job time : 45.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2004, 13:41:36 ; Search time 52 Seconds
(without alignments)
2325.583 Million cell updates/sec

Title: US-10-091-085-7

Perfect score: 2352

Sequence: 1 MATSWGTFVFLVVCVCSA.....ETGALGATPHLLQSLGISH 428

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq_29Jan04.*

- 1: Geneseq_1980s.*
- 2: Geneseq_1990s.*
- 3: Geneseq_2000s.*
- 4: Geneseq_2001s.*
- 5: Geneseq_2002s.*
- 6: Geneseq_2003as.*
- 7: Geneseq_2003bs.*
- 8: Geneseq_2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2252	100.0	428	3 AAY44850	Aay44850 Human CD3
2	2252	100.0	428	4 AAB72240	Aab72240 Human CD3
3	2235	99.2	428	3 AAY44849	Aay44849 Human CD3
4	2235	99.2	428	4 AAB72238	Aab72238 Human CD3
5	2235	99.2	428	4 AAB72243	Aab72243 Human CD3
6	2235	99.2	428	5 AAE19883	Aae19883 Human CD3
7	2089	92.8	405	3 AAY44851	Aay44851 Human CD3
8	2089	92.8	405	4 AAB72239	Aab72239 Human CD3
9	1817.5	80.7	465	5 AAE19884	Aae19884 Mouse CD3
10	1645	73.0	330	3 AAB53336	Aab53336 Human col
11	996	44.2	456	5 AAE19881	Aae19881 Human CD3
12	996	44.2	484	4 AAB72241	Aab72241 Human CD3
13	993	44.1	463	5 ABB06124	Abb06124 Human NS
14	993	44.1	467	5 ABO04657	Abj04657 Protein o
15	989	43.9	456	4 AAM93929	Aam93929 Human pol
16	986	43.8	446	5 AAB70458	Abj04658 Protein o
17	938	41.7	450	7 ADC14220	Adc14220 Human enz
18	894	39.7	462	4 AAU30882	Aau30882 Novel hum
19	816.5	36.3	471	4 AAB72242	Aab72242 Mature hu
20	693.5	30.8	461	4 ABB66213	Abb66213 Drosophil
21	693.5	30.8	464	4 ABB59611	Abb59611 Drosophil
22	605	26.9	476	3 AAY70912	Aay70912 Human CD3
23	605	26.9	476	3 AAY70889	Aay70889 Protein e
24	548	24.3	476	3 AAY70911	Aay70911 Human CD3
25	548	24.3	476	3 AAY70888	Aay70888 Protein e

ALIGNMENTS

RESULT 1

AAY44850

ID AAY44850 standard; protein; 428 AA.

XX

AC AAY44850;

AC

XX 18-MAY-2000 (first entry)

DT

XX Human CD39-L4 variant-ACR III mutant protein.

DB

XX

KW CD39-L4; human; apyrase; nucleotide diphosphatase; NDPase; variant;
KW ATP diphosphohydrolase; ATPase; adenosine diphosphate; ADP; treatment;
KW platelet aggregation; antithrombotic; thrombosis; myocardial infarction;
KW cerebral ischemia; angina; vascular graft; extracorporeal circulation;
KW molecular weight marker; nutritional supplement; tumour; prevention;
KW drug targeting; substitution mutation.

XX

OS Homo sapiens.

OS Synthetic.

XX

PH Key Location/Qualifiers

FT

FT Misc-difference 168 /note= "Wild type Asp substituted with Thr"

FT

FT Misc-difference 170 /note= "Wild type Ser substituted with Gln"

FT

FT Misc-difference 175 /note= "Wild type Leu substituted with Phe"

XX

PN WC200004041-A2.

XX

PD 27-JAN-2000.

XX

PF 16-JUL-1999; 99WO-US016180.

XX

PR 16-JUL-1998; 98US-00118205.

PR

PR 24-JUL-1998; 98US-00122449.

PR

PR 04-FEB-1999; 99US-00244444.

PR

PR 19-MAR-1999; 99US-00273447.

PR

PR 09-JUL-1999; 99US-00350836.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Ford J, Mulero J;

XX

XX WPI; 2000-182397/16.

DR

DR N-PSDS; AAZ50357.

XX

PT New nucleic acid encoding human CD39-like protein, useful for treating

PT and preventing thrombotic disease.

PS Claim 17; Fig 6; 125pp; English.

XX The present amino acid sequence is the CD39-L4 variant, designated as ACR
CC III mutant protein, an apyrase and/or nucleotide diphosphatase (NDPase).
CC It is isolated from the human foetal liver-spleen cDNA library,
CC b2HPLS20W. It is a soluble ATP diphosphohydrolase (ATPase) and is
CC involved in the hydrolysis of adenosine diphosphate (ADP), the agonist
CC that causes platelet aggregation. CD39-L4 protein has 30% and 80%
CC homology to human and murine CD39. It has platelet aggregation inhibition
CC and antithrombotic activity. CD39-L4 is used to treat or prevent
CC thrombosis, myocardial infarction, cerebral ischaemia and angina. It is
CC also used in vitro, to maintain vascular grafts or during extracorporeal
CC circulation, to hydrolyse NDP, as molecular weight markers and as
CC nutritional supplements. It is used to identify therapeutic agents that
CC bind and modulate CD39-L4. It is coupled to toxins for targeting drugs to
CC tumours or other cells that express CD39-L4

XX Sequence 428 AA;

Query Match 100.0%; Score 2252; DB 3; Length 428;

Best Local Similarity 100.0%; Pred. No. 9e-209;

Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATSGTFFPMLVSVCSAVSHRNQQTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT 60
DB 1 MATSGTFFPMLVSVCSAVSHRNQQTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT 60
QY 61 RIHVYTFVQKMPGQLPILGEVFDVSKPGLSAFVDPQKGAETVQGLLEVAKDSIPRSHW 120
DB 61 RIHVYTFVQKMPGQLPILGEVFDVSKPGLSAFVDPQKGAETVQGLLEVAKDSIPRSHW 120
QY 121 KKTIPVVKATAGLRLLPEHKAKALLPEVKEIFRKSPPFLVPKGSVIMTQDEGIFAMVTV 180
DB 121 KKTIPVVKATAGLRLLPEHKAKALLPEVKEIFRKSPPFLVPKGSVIMTQDEGIFAMVTV 180
QY 181 NFLTQLHGHROETVGTLDLGGASTQTITFLPOFEXTLEQTPRGYLTSEMFNSTYKLYTH 240
DB 181 NFLTQLHGHROETVGTLDLGGASTQTITFLPOFEXTLEQTPRGYLTSEMFNSTYKLYTH 240
QY 241 SYLPGKLAARLALGALTEGTDGTFRSACLPRWLEAEIFPGVKTYQYGNQGEVGF 300
DB 241 SYLPGKLAARLALGALTEGTDGTFRSACLPRWLEAEIFPGVKTYQYGNQGEVGF 300
QY 301 EPCYAEVLVRVVKGLHQPEVQSGSYAFSYVYDRAVTDMDIDYKGGILKVBDFPERKAR 360
DB 301 EPCYAEVLVRVVKGLHQPEVQSGSYAFSYVYDRAVTDMDIDYKGGILKVBDFPERKAR 360
QY 361 EVCNLENFTSGSPFLCWDLSYITALLKDGFGFADSTVLQLTQKVNNTETCWALCATPHL 420
DB 361 EVCNLENFTSGSPFLCWDLSYITALLKDGFGFADSTVLQLTQKVNNTETCWALCATPHL 420
QY 421 LQSLGISH 428
DB 421 LQSLGISH 428

RESULT 2

AAB72240

ID AAB72240 standard; protein; 428 AA.

XX AAB72240;

XX AAB72240;

DT 14-MAY-2001 (first entry)

XX Human CD39 like protein CD39-L4 variant ACRIII amino acid sequence.

XX Human CD39-like protein; apyrase; NDPase; platelet function inhibitor;
KW myocardial infarction; cerebral ischaemia; angina; arterial thrombosis;
KW cerebral artery thrombosis; platelet aggregation; inflammation;
KW apoptosis; autoimmune disorder; neurological disorder; mutant; mutein;
KW Alzheimer's disease; Parkinson's disease; cancer; CD39-L4.

XX Homo sapiens.

OS WC200110205-A1.

PN 15-FEB-2001.

XX 03-AUG-2000; 2000WO-US021790.

XX 09-AUG-1999; 99US-00370265.

XX 11-JAN-2000; 2000US-00481238.

XX 25-APR-2000; 2000US-00557800.

XX 28-MAY-2000; 2000US-00582231.

XX 30-JUN-2000; 2000US-00608285.

XX (HYSE-) HYSEQ INC.

XX Ford J, Mulero JJ, Yeung G;

XX WPI; 2001-147489/15.

XX N-PSDB; AAF63385.

XX Polynucleotides encoding human CD39-like polypeptides, with apyrase

XX and/or NDPase activity, which are useful in the treatment of pathological

XX conditions caused by thrombosis (e.g. myocardial infarction) and

XX inflammatory disorders.

XX Claim 19; Fig 6; 203pp; English.

XX This invention relates to polynucleotides encoding human CD39-like

XX polypeptides with apyrase and/or NDPase activity. The polypeptides having

XX ATPase, including NDPase, activity are useful for inhibiting platelet

XX function and can therefore be used in the prophylaxis or treatment of

XX pathological conditions caused by or involving thrombosis or excessive

XX coagulation or excessive platelet aggregation, such as myocardial

XX infarction, cerebral ischaemia, angina, arterial thrombosis, cerebral

XX artery thrombosis or intracardiac thrombosis, and conditions associated

XX with venous thrombosis. CD39-L4 and CD39-L2 polypeptides are useful in

XX modulating disease states (including platelet aggregation, inflammation

XX and apoptosis) associated with ADP or other purinergic signalling by

XX reducing the levels of NDPs. The polypeptides are also useful for

XX prophylaxis or treatment of inflammation related disorders, such as

XX disorders involving sepsis or systemic inflammatory response syndrome or

XX SIRS (and associated conditions such as fever, tachycardia, tachypnea,

XX cytokine overstimulation); autoimmune disorders such as thrombosis,

XX atherosclerosis, acute pancreatitis, dermatitis, including psoriasis,

XX cirrhosis, reperfusion injury, asthma, multiple sclerosis, arthritis,

XX neurological disorders including neurodegenerative diseases, epilepsy,

XX depression, Alzheimer's disease, Parkinson's disease, Huntington's

XX disease, and amyotrophic lateral sclerosis; and cancer. The present

XX sequence represents human CD39 like protein CD39-L4 variant ACRIII

XX Sequence 428 AA;

Query Match 100.0%; Score 2252; DB 4; Length 428;

Best Local Similarity 100.0%; Pred. No. 9e-209;

Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATSGTFFPMLVSVCSAVSHRNQQTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT 60

DB 1 MATSGTFFPMLVSVCSAVSHRNQQTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT 60

QY 61 RIHVYTFVQKMPGQLPILGEVFDVSKPGLSAFVDPQKGAETVQGLLEVAKDSIPRSHW 120

DB 61 RIHVYTFVQKMPGQLPILGEVFDVSKPGLSAFVDPQKGAETVQGLLEVAKDSIPRSHW 120

QY 121 KKTIPVVKATAGLRLLPEHKAKALLPEVKEIFRKSPPFLVPKGSVIMTQDEGIFAMVTV 180

DB 121 KKTIPVVKATAGLRLLPEHKAKALLPEVKEIFRKSPPFLVPKGSVIMTQDEGIFAMVTV 180

QY 181 NFLTQLHGHROETVGTLDLGGASTQTITFLPOFEXTLEQTPRGYLTSEMFNSTYKLYTH 240

DB 181 NFLTQLHGHROETVGTLDLGGASTQTITFLPOFEXTLEQTPRGYLTSEMFNSTYKLYTH 240

QY 241 SYLGFLKAARLATLGALETGDTGHTFRSACLPRLAEAFWIFGGVKYQYGGNQEVEGVF 300
Db 241 SYLGFLKAARLATLGALETGDTGHTFRSACLPRLAEAFWIFGGVKYQYGGNQEVEGVF 300
QY 301 EPCYAEVLVRVGRKHLQHPREVORGSFYAFSYVDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 301 EPCYAEVLVRVGRKHLQHPREVORGSFYAFSYVDRAVDTMDIDYKGGILKVEDFERKAR 360
QY 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTCKVNNIETGWALGATPHL 420
Db 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTCKVNNIETGWALGATPHL 420
QY 421 LQSLGISH 428
Db 421 LQSLGISH 428

RESULT 3
AA44849
ID AAY44849 standard; protein; 428 AA.
AC AAY44849;
XX 18-MAY-2000 (first entry)
DT Human CD39-L4 protein.
DE
XX CD39-L4; human; apyrase; nucleotide diphosphatase; NDPase;
KW ATP diphosphohydrolase; ATPase; adenosine diphosphate; ADP; treatment;
KW platelet aggregation; antithrombotic; thrombosis; myocardial infarction;
KW cerebral ischaemia; angina; vascular graft; extracorporeal circulation;
KW molecular weight marker; nutritional supplement; tumour; prevention;
KW drug targeting; Apyrase Conserved Region; ACR.
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT Peptide 1..22 /label= Leader_peptide
FT Protein 23..428 /label= Mature_human_CD39-L4_protein
FT /note= "Homologous to human and murine CD39"
FT Binding-site 54..58 /label= ATP_Binding_region
FT Region 129..134 /label= Apyrase_Conserved_Region
FT Region 169..173 /label= Apyrase_Conserved_Region
FT Region 199..206 /note= "Conserved motif in ATPases"
XX WO200004041-A2.
PN
XX
XX 27-JAN-2000.
XX
PF 16-JUL-1999; 99WO-US016180.
XX
PR 16-JUL-1998; 98US-00118205.
PR 24-JUL-1998; 98US-00124439.
PR 04-FEB-1999; 99US-00244444.
PR 13-MAR-1999; 99US-00273447.
PR 09-JUL-1999; 99US-00350836.
XX
XX (HYSE-) HYSEQ INC.
XX Ford J, Mulero J;
XX WPI; 2000-182397/16.
DR N-PSDB; AAZ50356; AAZ50359.
XX New nucleic acid encoding human CD39-like protein, useful for treating
PT and preventing thrombotic disease.
XX

XX Claim 15; Fig 2; 125pp; English.
PS The present amino acid sequence is the CD39-L4 protein, an apyrase and/or
XX nucleotide diphosphatase (NDPase). It is isolated from the human foetal
CC liver-spleen cDNA library, b2HPLS20W. It is a soluble ATP
CC diphosphohydrolase (ATPDase) and is involved in the hydrolysis of
CC adenosine diphosphate (ADP), the agonist that causes platelet
CC aggregation. CD39-L4 protein has 30% and 80% homology to human and murine
CC CD39. It has platelet aggregation inhibition and antithrombotic activity.
CC CD39-L4 is used to treat or prevent thrombosis, myocardial infarction,
CC cerebral ischaemia and angina. It is also used in vitro, to maintain
CC vascular grafts or during extracorporeal circulation, to hydrolyse NDP,
CC as molecular weight markers and as nutritional supplements. It is used to
CC identify therapeutic agents that bind and modulate CD39-L4. It is coupled
CC to toxins for targeting drugs to tumours or other cells that express CD39
XX -L4
XX
SQ Sequence 428 AA;
Query Match 99.2%; Score 2235; DB 3; Length 428;
Best Local Similarity 99.3%; Pred. No. 4e-207;
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 MATSWGTVFFMLVWSCVSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSWGTVFFMLVWSCVSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
QY 61 RIHVYTFVQKMPGQLPILGEVFDVSKPGLSAFVDPKQGAFTVQGLLEVAKDSIPRSHW 120
Db 61 RIHVYTFVQKMPGQLPILGEVFDVSKPGLSAFVDPKQGAFTVQGLLEVAKDSIPRSHW 120
QY 121 KKTVPVLKATAGLRLLPEKAKALLFEVKEIFPKSPFLVPKGSVIMTQDREGIFAWVTV 180
Db 121 KKTVPVLKATAGLRLLPEKAKALLFEVKEIFPKSPFLVPKGSVIMTQDREGIFAWVTV 180
QY 181 NFLTQQLHGHROETVGTLDLGGASTQITFLPQFETKLTGOTPRGYLTSPFMFNSTYKLYTH 240
Db 181 NFLTQQLHGHROETVGTLDLGGASTQITFLPQFETKLTGOTPRGYLTSPFMFNSTYKLYTH 240
QY 241 SYLGFLKAARLATLGALETGDTGHTFRSACLPRLAEAFWIFGGVKYQYGGNQEVEGVF 300
Db 241 SYLGFLKAARLATLGALETGDTGHTFRSACLPRLAEAFWIFGGVKYQYGGNQEVEGVF 300
QY 301 EPCYAEVLVRVGRKHLQHPREVORGSFYAFSYVDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 301 EPCYAEVLVRVGRKHLQHPREVORGSFYAFSYVDRAVDTMDIDYKGGILKVEDFERKAR 360
QY 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTCKVNNIETGWALGATPHL 420
Db 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTCKVNNIETGWALGATPHL 420
QY 421 LQSLGISH 428
Db 421 LQSLGISH 428

RESULT 4
AAB72238
ID AAB72238 standard; protein; 428 AA.
XX
AC AAB72238;
XX 14-MAY-2001 (first entry)
DT Human CD39 like protein CD39-L4 amino acid sequence.
DE
XX Human CD39-like protein; apyrase; NDPase; platelet function inhibitor;
KW myocardial infarction; cerebral ischaemia; angina; arterial thrombosis;
KW cerebral artery thrombosis; platelet aggregation; inflammation;
KW apoptosis; autoimmune disorder; neurological disorder;
KW Alzheimer's disease; Parkinson's disease; cancer; CD39-L4.
XX

OS Homo sapiens.
 XX WO200110205-A1.
 XX
 XX PD 15-FEB-2001.
 XX
 XX PF 09-AUG-2000; 2000WO-US021790.
 XX
 XX PR 09-AUG-1999; 99US-00370265.
 XX PR 11-JAN-2000; 2000US-00481238.
 XX PR 25-APR-2000; 2000US-00557800.
 XX PR 26-MAY-2000; 2000US-00583231.
 XX PR 30-JUN-2000; 2000US-00608285.
 XX
 XX PA (HYSE-) HYSEQ INC.
 XX
 XX PI Ford J, Mulero JJ, Yeung G;
 XX
 XX DR WPI; 2001-147489/15.
 XX DR N-PSDB; AAF63383.
 XX
 XX PT Polynucleotides encoding human CD39-like polypeptides, with apyrase
 XX and/or NDase activity, which are useful in the treatment of pathological
 XX conditions caused by thrombosis (e.g. myocardial infarction) and
 XX inflammatory disorders.
 XX
 XX PS Claim 15; Fig 2; 203pp; English.
 XX
 XX CC This invention relates to polynucleotides encoding human CD39-like
 XX polypeptides with apyrase and/or NDase activity. The polypeptides having
 XX ATPase, including NDase, activity are useful for inhibiting platelet
 XX function and can therefore be used in the prophylaxis or treatment of
 XX pathological conditions caused by or involving thrombosis or excessive
 XX coagulation or excessive platelet aggregation, such as myocardial
 XX infarction, cerebral ischaemia, angina, arterial thrombosis, cerebral
 XX artery thrombosis or intracardiac thrombosis, and conditions associated
 XX with venous thrombosis. CD39-L4 and CD39-L2 polypeptides are useful in
 XX modulating disease states (including platelet aggregation, inflammation
 XX and apoptosis) associated with ADP or other purinergic signalling by
 XX reducing the levels of NDps. The polypeptides are also useful for
 XX prophylaxis or treatment of inflammation related disorders, such as
 XX disorders involving sepsis or systemic inflammatory response syndrome or
 XX SIRS (and associated conditions such as fever, tachycardia, tachypnea,
 XX cytokine overstimulation); autoimmune disorders such as thrombosis,
 XX atherosclerosis, acute pancreatitis, dermatitis, including psoriasis,
 XX cirrhosis, reperfusion injury, asthma, multiple sclerosis, arthritis;
 XX neurological disorders including neurodegenerative diseases, epilepsy,
 XX depression, Alzheimer's disease, Parkinson's disease, Huntington's
 XX disease, and amyotrophic lateral sclerosis; and cancer. The present
 XX sequence represents human CD39 like protein CD39-L4
 XX
 XX SQ Sequence 428 AA;
 XX
 XX Query Match 99.2%; Score 2235; DB 4; Length 428;
 XX Best Local Similarity 99.3%; Pred. No. 4e-207;
 XX Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 XX
 XX QY 1 MATSGTVPFVLVSVCSAVSHENQQTWTFGIFLSSMCPINVSASTLYGIMEDAGSTGT 60
 XX DB 1 MATSGTVPFVLVSVCSAVSHENQQTWTFGIFLSSMCPINVSASTLYGIMEDAGSTGT 60
 XX
 XX QY 61 RIHYVTFVQKMPGQLPILGEGVFDSPKGLSAFVDQPKQGAETVQGLLEKADSIKPSHW 120
 XX DB 61 RIHYVTFVQKMPGQLPILGEGVFDSPKGLSAFVDQPKQGAETVQGLLEKADSIKPSHW 120
 XX
 XX QY 121 KKTPLVKATAGRLLPKHKALLFEVKEIFRKSPLVPGKSVIMTGDGIFAWTV 180
 XX DB 121 KKTPLVKATAGRLLPKHKALLFEVKEIFRKSPLVPGKSVIMTGDGIFAWTV 180
 XX
 XX QY 181 NFLTQGLHCHQETVGTLDGASTQITFLPQFKTLEQTPRGYLTSEFNFSTYKLYTH 240
 XX DB 181 NFLTQGLHCHQETVGTLDGASTQITFLPQFKTLEQTPRGYLTSEFNFSTYKLYTH 240

241 SYLGFLKAARLATIATLGALETEGTGHTFRSACLPRWLEAEMIFGGVKYQYGNQGEVGF 300
 |||||
 241 SYLGFLKAARLATIATLGALETEGTGHTFRSACLPRWLEAEMIFGGVKYQYGNQGEVGF 300
 |||||
 301 EPCYAEVLVRVGRKLIHQPEEVQSGSFYAESYYDRAVDMDIDYEKGGLIKVEDFFRKAR 360
 |||||
 301 EPCYAEVLVRVGRKLIHQPEEVQSGSFYAESYYDRAVDMDIDYEKGGLIKVEDFFRKAR 360
 |||||
 361 EVCNLENFTSGSPFLCWDLSYITALLKDGFGFADSTVLQLTCKVNNIFTGALGATPHL 420
 |||||
 361 EVCNLENFTSGSPFLCWDLSYITALLKDGFGFADSTVLQLTCKVNNIFTGALGATPHL 420
 |||||
 421 LQSLGISH 428
 |||||
 421 LQSLGISH 428
 |||||

RESULT 5
 AAB72243
 ID AAB72243 standard; protein; 428 AA.
 XX
 XX AC AAB72243;
 XX
 XX DT 14-MAY-2001 (first entry)
 XX
 XX DE Human CD39 like protein CD39-L4 amino acid sequence.
 XX
 XX KW Human CD39-like protein; apyrase; NDase; platelet function inhibitor;
 XX myocardial infarction; cerebral ischaemia; angina; arterial thrombosis;
 XX cerebral artery thrombosis; platelet aggregation; inflammation;
 XX apoptosis; autoimmune disorder; neurological disorder;
 XX Alzheimer's disease; Parkinson's disease; cancer; CD39-L4.
 XX
 XX OS Homo sapiens.
 XX
 XX PN WO200110205-A1.
 XX
 XX PD 15-FEB-2001.
 XX
 XX PF 09-AUG-2000; 2000WO-US021790.
 XX
 XX PR 09-AUG-1999; 99US-00370265.
 XX PR 11-JAN-2000; 2000US-00481238.
 XX PR 25-APR-2000; 2000US-00557800.
 XX PR 26-MAY-2000; 2000US-00583231.
 XX PR 30-JUN-2000; 2000US-00608285.
 XX
 XX PA (HYSE-) HYSEQ INC.
 XX
 XX PI Ford J, Mulero JJ, Yeung G;
 XX
 XX DR WPI; 2001-147489/15.
 XX DR N-PSDB; AAF63402.
 XX
 XX PT Polynucleotides encoding human CD39-like polypeptides, with apyrase
 XX and/or NDase activity, which are useful in the treatment of pathological
 XX conditions caused by thrombosis (e.g. myocardial infarction) and
 XX inflammatory disorders.
 XX
 XX PS Example 9; Page 142-144; 203pp; English.
 XX
 XX CC This invention relates to polynucleotides encoding human CD39-like
 XX polypeptides with apyrase and/or NDase activity. The polypeptides having
 XX ATPase, including NDase, activity are useful for inhibiting platelet
 XX function and can therefore be used in the prophylaxis or treatment of
 XX pathological conditions caused by or involving thrombosis or excessive
 XX coagulation or excessive platelet aggregation, such as myocardial
 XX infarction, cerebral ischaemia, angina, arterial thrombosis, cerebral
 XX artery thrombosis or intracardiac thrombosis, and conditions associated
 XX with venous thrombosis. CD39-L4 and CD39-L2 polypeptides are useful in
 XX modulating disease states (including platelet aggregation, inflammation
 XX and apoptosis) associated with ADP or other purinergic signalling by
 XX reducing the levels of NDps. The polypeptides are also useful for

CC prophylaxis or treatment of inflammation related disorders, such as
 CC disorders involving sepsis or systemic inflammatory response syndrome or
 CC SIRS (and associated conditions such as fever, tachycardia, tachypnea,
 CC cytokine overstimulation); autoimmune disorders such as thrombosis,
 CC atherosclerosis, acute pancreatitis, dermatitis, including psoriasis,
 CC cirrhosis, reperfusion injury, asthma, multiple sclerosis, arthritis;
 CC neurological disorders including neurodegenerative diseases, epilepsy,
 CC depression, Alzheimer's disease, Parkinson's disease, Huntington's
 CC disease, and amyotrophic lateral sclerosis; and cancer. The present
 CC sequence represents the CD39 like protein CD39-L4 amino acid sequence
 XX
 SQ Sequence 428 AA;

Query Match 99.2%; Score 2235; DB 4; Length 428;
 Best Local Similarity 99.3%; Pred. No. 4e-207;
 Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MATSGTFFPMLVWSCVCSAUSHRNQOTWFGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
 Dd 1 MATSGTFFPMLVWSCVCSAUSHRNQOTWFGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy 61 RIHVYTFVQKMPGQLPILGEVDSVKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120
 Dd 61 RIHVYTFVQKMPGQLPILGEVDSVKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120

Qy 121 KKTPTVLKATAGLRLLPEHKAKALLFEVKEIFPKSPFLVPKGSVIMTQDEGIFAWVTV 180
 Dd 121 KKTPTVLKATAGLRLLPEHKAKALLFEVKEIFPKSPFLVPKGSVIMTQDEGIFAWVTV 180

Qy 181 NPLTQGLHGHROETVGTLDLGASTQITFLPQFKTLQTPRGYLTSTFEMFNSTYKLYTH 240
 Dd 181 NPLTQGLHGHROETVGTLDLGASTQITFLPQFKTLQTPRGYLTSTFEMFNSTYKLYTH 240

Qy 241 SYLGFGLKAARLALGALTEGTDGHTFRSACLPRWLAEMIFPGVKYQYQGNQGEVGF 300
 Dd 241 SYLGFGLKAARLALGALTEGTDGHTFRSACLPRWLAEMIFPGVKYQYQGNQGEVGF 300

Qy 301 EPCYAEVLVRVGRKLPQEEVQSGSFYAFSYYYDRAVDTMDIDKGGILKVEDEKAR 360
 Dd 301 EPCYAEVLVRVGRKLPQEEVQSGSFYAFSYYYDRAVDTMDIDKGGILKVEDEKAR 360

Qy 361 EVCNDLENFTSGSPFLCDLSYITALLKDGFGFADSTVLQTKYNNIETQWALGATPHL 420
 Dd 361 EVCNDLENFTSGSPFLCDLSYITALLKDGFGFADSTVLQTKYNNIETQWALGATPHL 420

Qy 421 LQSLGISH 428
 Dd 421 LQSLGISH 428

RESULT 6
 AAE19883
 ID AAE19883 standard; protein; 428 AA.

XX AAE19883;

XX 18-JUN-2002 (first entry)

XX Human CD39L4 protein.

XX Human; CD-39-like protein; CD39L4 protein; therapy; immune deficiency;
 KW autoimmune disorder; multiple sclerosis; systemic lupus erythematosus;
 KW rheumatoid arthritis; autoimmune thyroiditis; allergic reaction; asthma;
 KW insulin dependent diabetes mellitus; periodontal disease; osteoporosis;
 KW osteoarthritis; wound healing; tissue repair; Alzheimer's disease; ulcer;
 KW Parkinson's disease; amyotrophic lateral sclerosis; Huntington's disease;
 KW nervous system disease; nerve injury; ischaemia-reperfusion injury;
 KW endotoxin lethality; arthritis; nephritis; inflammatory bowel disease;
 KW Crohn's disease; virucide; antibacterial; antifungal; neuroprotective;
 KW dermatological; immunosuppressive; vulnary; neurotropic; anticonvulsant;
 KW antiinflammatory; nephrotropic; gastrointestinal; vasotropic.

OS Homo sapiens.

XX US6350447-B1.
 PN
 XX 26-FEB-2002.
 PD
 XX 29-JAN-1999; 99US-00240639.
 PF
 XX 29-JAN-1999; 99US-00240639.
 PR
 XX (HYSB-) HYSEQ INC.
 PA
 XX Chadwick BP, Frischauf A;
 PI
 XX WPI; 2002-215262/27.
 DR
 XX N-PSDB; AAD31695.
 DR
 XX
 PT An isolated polypeptide with phosphohydrolase activity, designated
 PT CD39L2, useful to identify other proteins with which binding occurs or
 PT identify inhibitors and for treatment of, e.g., Alzheimer's, multiple
 PT sclerosis and osteoporosis.
 XX
 PS Example; Fig 7; 101pp; English.

CC The present invention relates to novel proteins with phosphohydrolase
 CC activity, designated CD-39-like (CD39L) proteins and polynucleotides
 CC encoding such proteins. CD39L proteins are useful to treat infectious
 CC diseases caused by viral, bacterial, fungal or other infection that may
 CC be treatable with CD39L. They are useful in the treatment of various
 CC immune deficiencies and disorders, autoimmune disorders such as multiple
 CC sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune
 CC thyroiditis and insulin dependent diabetes mellitus, allergic reactions
 CC and conditions such as asthma and other respiratory problems, periodontal
 CC disease, osteoporosis, osteoarthritis and other tooth repair processes.
 CC They may have utility in compositions used for bone, cartilage, tendon,
 CC ligament and/or nerve tissue growth or regeneration as well as for wound
 CC healing and tissue repair and replacement and in the treatment of burns,
 CC incisions and ulcers. CD39L proteins may also be useful for proliferation
 CC of neural cells and for regeneration of nerve and brain tissue, i.e. for
 CC the treatment of central nervous system diseases such as Alzheimer's
 CC disease, Parkinson's disease, amyotrophic lateral sclerosis, Huntington's
 CC disease, peripheral nervous system diseases peripheral nerve injuries,
 CC peripheral neuropathy and localised neuropathies. They are also used to
 CC treat mechanical and traumatic disorders which involve degeneration,
 CC death or trauma to neural cells or nerve tissue. CD39L proteins of the
 CC invention are also useful to promote better or faster closure of non-
 CC healing wounds, including pressure ulcers, ulcers associated with
 CC vascular insufficiency and surgical and traumatic wounds. They also
 CC exhibit anti-inflammatory activity and may be used to treat inflammatory
 CC conditions including chronic or acute conditions), including ischaemia-
 CC reperfusion injury, endotoxin lethality, arthritis, nephritis, cytokine
 CC or chemokine-induced lung injury, inflammatory bowel disease or Crohn's
 CC disease. The present sequence is human CD39L4 protein

XX SQ Sequence 428 AA;

Query Match 99.2%; Score 2235; DB 5; Length 428;
 Best Local Similarity 99.3%; Pred. No. 4e-207;
 Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MATSGTFFPMLVWSCVCSAUSHRNQOTWFGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
 Dd 1 MATSGTFFPMLVWSCVCSAUSHRNQOTWFGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy 61 RIHVYTFVQKMPGQLPILGEVDSVKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120
 Dd 61 RIHVYTFVQKMPGQLPILGEVDSVKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120

Qy 121 KKTPTVLKATAGLRLLPEHKAKALLFEVKEIFPKSPFLVPKGSVIMTQDEGIFAWVTV 180
 Dd 121 KKTPTVLKATAGLRLLPEHKAKALLFEVKEIFPKSPFLVPKGSVIMTQDEGIFAWVTV 180

Qy 181 NPLTQGLHGHROETVGTLDLGASTQITFLPQFKTLQTPRGYLTSTFEMFNSTYKLYTH 240
 Dd 181 NPLTQGLHGHROETVGTLDLGASTQITFLPQFKTLQTPRGYLTSTFEMFNSTYKLYTH 240

DB 181 NLTQLHGRQETVGTLDLGGASTQITFLPQFEKTLBQTPRGYLTSFEMFNSTYKLYTH 240
 QY 241 SYLGFLKAARLALCALTEGTDGHTFRSACLPRMLEAEIFGGVKYQYGNQGEVGF 300
 DB 241 SYLGFLKAARLALCALTEGTDGHTFRSACLPRMLEAEIFGGVKYQYGNQGEVGF 300
 QY 301 EPCVAEVLVRVGRKLGHPSEVQVGSFYAFSYIYDRAVDTMDIDYKGGILKVEDFERKAR 360
 DB 301 EPCVAEVLVRVGRKLGHPSEVQVGSFYAFSYIYDRAVDTMDIDYKGGILKVEDFERKAR 360
 QY 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQTKKVNNTETGALCATPHL 420
 DB 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQTKKVNNTETGALCATPHL 420
 QY 421 LOSLGISH 428
 DB 421 LOSLGISH 428

RESULT 7
 AAY44851
 ID AAY44851 standard; protein; 405 AA.
 AC AAY44851;
 XX
 DT 18-MAY-2000 (first entry)
 DE Human CD39-L66 protein.
 XX
 KW CD39-L66; human; CD39-L66; apyrase; nucleotide diphosphatase; NDPase;
 KW ATP Diphosphohydrolase; ATPDase; adenosine diphosphate; ADP; treatment;
 KW platelet aggregation; antithrombotic; thrombosis; myocardial infarction;
 KW cerebral ischaemia; angina; vascular graft; extracorporeal circulation;
 KW molecular weight marker; nutritional supplement; tumour; prevention;
 KW drug targeting; splice variant.
 XX
 OS Homo sapiens.
 XX
 PN WO200004041-A2.
 XX
 PD 27-JAN-2000.
 XX
 PF 16-JUL-1999; 99WO-US016180.
 XX
 PR 16-JUL-1998; 98US-00118205.
 PR 24-JUL-1998; 98US-00122449.
 PR 04-FEB-1999; 98US-00244444.
 PR 19-MAR-1999; 98US-00273447.
 PR 09-JUL-1999; 99US-00350836.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Ford J, Mulero J;
 XX
 DR WPI: 2000-182397/16.
 DR N-PSDB; AA250358.
 XX

PT New nucleic acid encoding human CD39-like protein, useful for treating
 PT and preventing thrombotic disease.
 XX
 PS Claim 15; Page 124-125; 125pp; English.
 XX
 CC The present amino acid sequence is the CD39-L66 protein, a splice variant
 CC of the CD39-L4 protein. It is an apyrase and/or nucleotide diphosphatase
 CC (NDPase), isolated from the human foetal liver-spleen cDNA library,
 CC b2HPLS20W. It is a soluble ATP Diphosphohydrolase (ATPDase) and is
 CC involved in the hydrolysis of adenosine diphosphate (ADP), the agonist
 CC that causes platelet aggregation. CD39-L4 protein has 30% and 80%
 CC homology to human and murine CD39. It has platelet aggregation inhibition
 CC and antithrombotic activity. CD39-L4 is used to treat or prevent
 CC thrombosis, myocardial infarction, cerebral ischaemia and angina. It is
 CC also used in vitro, to maintain vascular grafts or during extracorporeal
 CC circulation, to hydrolyse NDP, as molecular weight markers and as

CC nutritional supplements. It is used to identify therapeutic agents that
 CC bind and modulate CD39-L4. It is coupled to toxins for targeting drugs to
 CC tumours or other cells that express CD39-L4
 XX
 SQ Sequence 405 AA;

Query Match 92.8%; Score 2089; DB 3; Length 405;
 Best Local Similarity 99.2%; Pred. No. 4.9e-193;
 Matches 397; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MATSWGTVFPMVLVSCVSAVSHRNQOTWFEIGIFLSMPCPINSASTLYGIMFDAGSTGT 60
 DB 1 MATSWGTVFPMVLVSCVSAVSHRNQOTWFEIGIFLSMPCPINSASTLYGIMFDAGSTGT 60
 QY 61 RIHYTVFQMPQQLPILEGVFDVSKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120
 DB 61 RIHYTVFQMPQQLPILEGVFDVSKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120
 QY 121 KKTPTVLKATAGRLLLPEHAKALLPEVKIIFRKSPLVPKGSVIMTGDGEGIFAWTV 180
 DB 121 KKTPTVLKATAGRLLLPEHAKALLPEVKIIFRKSPLVPKGSVIMTGDGEGIFAWTV 180
 QY 181 NLTQLHGRQETVGTLDLGGASTQITFLPQFEKTLBQTPRGYLTSFEMFNSTYKLYTH 240
 DB 181 NLTQLHGRQETVGTLDLGGASTQITFLPQFEKTLBQTPRGYLTSFEMFNSTYKLYTH 240
 QY 241 SYLGFLKAARLALCALTEGTDGHTFRSACLPRMLEAEIFGGVKYQYGNQGEVGF 300
 DB 241 SYLGFLKAARLALCALTEGTDGHTFRSACLPRMLEAEIFGGVKYQYGNQGEVGF 300
 QY 301 EPCVAEVLVRVGRKLGHPSEVQVGSFYAFSYIYDRAVDTMDIDYKGGILKVEDFERKAR 360
 DB 301 EPCVAEVLVRVGRKLGHPSEVQVGSFYAFSYIYDRAVDTMDIDYKGGILKVEDFERKAR 360
 QY 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQ 400
 DB 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQ 400

RESULT 8
 AAB72239
 ID AAB72239 standard; protein; 405 AA.
 AC AAB72239;
 XX
 DT 14-MAY-2001 (first entry)
 DE Human CD39 like protein CD39-L4 amino acid sequence.
 XX
 KW Human CD39-like protein; apyrase; NDPase; platelet function inhibitor;
 KW myocardial infarction; cerebral ischaemia; angina; arterial thrombosis;
 KW cerebral artery thrombosis; platelet aggregation; inflammation;
 KW apoptosis; autoimmune disorder; neurological disorder;
 KW Alzheimer's disease; Parkinson's disease; cancer; CD39-L4.
 XX
 OS Homo sapiens.
 XX
 PN WO200110205-A1.
 XX
 PD 15-FEB-2001.
 XX
 PF 09-AUG-2000; 2000WO-US021790.
 XX
 PR 09-AUG-1999; 99US-00370265.
 PR 11-JAN-2000; 2000US-00481238.
 PR 25-APR-2000; 2000US-00557800.
 PR 26-MAY-2000; 2000US-00583231.
 PR 30-JUN-2000; 2000US-00608285.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Ford J, Mulero JJ, Yeung G;
 XX

DR WPI; 2001-147489/15.
DR N-PSDB; AAF63384.
XX Polynucleotides encoding human CD39-like polypeptides, with apyrase
PT and/or NTPase activity, which are useful in the treatment of pathological
PT conditions caused by thrombosis (e.g. myocardial infarction) and
PT inflammatory disorders.
PS Claim 17; Page 157-158; 203pp; English.
XX This invention relates to polynucleotides encoding human CD39-like
CC polypeptides with apyrase and/or NTPase activity. The polypeptides having
CC ATPase, including ADPase, activity are useful for inhibiting platelet
CC function and can therefore be used in the prophylaxis or treatment of
CC pathological conditions caused by or involving thrombosis or excessive
CC coagulation or excessive platelet aggregation, such as myocardial
CC infarction, cerebral ischaemia, angina, arterial thrombosis, cerebral
CC artery thrombosis or intracardiac thrombosis, and conditions associated
CC with venous thrombosis. CD39-L4 and CD39-L2 polypeptides are useful in
CC modulating disease states (including platelet aggregation, inflammation
CC and apoptosis) associated with ADP or other purinergic signalling by
CC reducing the levels of NDPs. The polypeptides are also useful for
CC prophylaxis or treatment of inflammation related disorders, such as
CC disorders involving sepsis or systemic inflammatory response syndrome or
CC SIRS (and associated conditions such as fever, tachycardia, tachypnea,
CC cytokine overstimulation); autoimmune disorders such as thrombosis,
CC atherosclerosis, acute pancreatitis, dermatitis, including psoriasis,
CC cirrhosis, reperfusion injury, asthma, multiple sclerosis, arthritis,
CC neurological disorders including neurodegenerative diseases, epilepsy,
CC depression, Alzheimer's disease, Parkinson's disease, Huntington's
CC disease, and amyotrophic lateral sclerosis; and cancer. The present
CC sequence represents human CD39 like protein CD39-L4
XX
XX Sequence 405 AA;

Query Match 92.8%; Score 2089; DB 4; Length 405;
Best Local Similarity 99.2%; Pred. No. 4.9e-193;
Matches 397; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 MATSWGTVPFVLMVSVCSAVSHRNQOTWFGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
DB 1 MATSWGTVPFVLMVSVCSAVSHRNQOTWFGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
QY 61 RIHVYTFVQMPQQLPILGEVPSVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
DB 61 RIHVYTFVQMPQQLPILGEVPSVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
QY 121 KKTPTVVLKATAGLRLPEKAKALLPEVKEIPKSPFLVPKGSVSIWTCQDGI FAWVTV 180
DB 121 KKTPTVVLKATAGLRLPEKAKALLPEVKEIPKSPFLVPKGSVSIWTCQDGI FAWVTV 180
QY 181 NFLTGLQHGRQETVGTGLDGGASTQTITLPQFEKLTQTPRGYLTSPFEMFNSTYKLYTH 240
DB 181 NFLTGLQHGRQETVGTGLDGGASTQTITLPQFEKLTQTPRGYLTSPFEMFNSTYKLYTH 240
QY 241 SYLGFGLKAARLATGALTEGDTGHTFRSACLPRWLEAEWIFGGVKYQYCGNQGEVGF 300
DB 241 SYLGFGLKAARLATGALTEGDTGHTFRSACLPRWLEAEWIFGGVKYQYCGNQGEVGF 300
QY 301 EPCYAEVLVRVGRKLHQPEEVORGSFYAFSYVDRAVDMDIDYKGGILKVEDFERKAR 360
DB 301 EPCYAEVLVRVGRKLHQPEEVORGSFYAFSYVDRAVDMDIDYKGGILKVEDFERKAR 360
QY 361 EVCNDLENFTSSPFLCMDLSTYITALLKDGQFGFADSTVLQ 400
DB 361 EVCNDLENFTSSPFLCMDLSTYITALLKDGQFGFADSTVLQ 400

RESULT 9
AAE19884
ID AAE19884 standard; protein; 465 AA.
XX
AC AAE19884;

XX DT 18-JUN-2002 (first entry)
DE Mouse CD39L4 protein.
XX
XX Mouse; CD-39-like protein; CD39L4 protein; therapy; immune deficiency;
KW autoimmune disorder; multiple sclerosis; systemic lupus erythematosus;
KW rheumatoid arthritis; autoimmune thyroiditis; allergic reaction; asthma;
KW insulin dependent diabetes mellitus; periodontal disease; osteoporosis;
KW osteoarthritis; wound healing; tissue repair; Alzheimer's disease; ulcer;
KW Parkinson's disease; amyotrophic lateral sclerosis; Huntington's disease;
KW nervous system disease; nerve injury; ischaemia-reperfusion injury;
KW endotoxin lethality; arthritis; nephritis; inflammatory bowel disease;
KW Crohn's disease; virucide; antibacterial; antifungal; neuroprotective;
KW dermatological; immunosuppressive; vulnary; neurotropic; anticonvulsant;
KW antiinflammatory; nephrotropic; gastrointestinal; vasotropic; NTPase;
KW nucleotide-triphosphatase; enzyme.
XX Mus musculus.
XX US6350447-B1.
PN 26-FEB-2002.
XX 29-JAN-1999; 99US-00240639.
XX 29-JAN-1999; 99US-00240639.
XX (HYSE-) HYSEQ INC.
PI Chadwick BP, Frischaut A;
DR WPI; 2002-215262/27.
DR N-PSDB; AAD31696.
XX
PT An isolated polypeptide with phosphohydrolase activity, designated
PT CD39L2, useful to identify other proteins with which binding occurs or
PT identify inhibitors and for treatment of, e.g., Alzheimer's, multiple
PT sclerosis and osteoporosis.
XX Example; Fig 1; 101pp; English.
XX The present invention relates to novel proteins with phosphohydrolase
CC activity, designated CD-39-like (CD39L) proteins and polynucleotides
CC encoding such proteins. CD39L proteins are useful to treat infectious
CC diseases caused by viral, bacterial, fungal or other infection that may
CC be treatable with CD39L. They are useful in the treatment of various
CC immune deficiencies and disorders, autoimmune disorders such as multiple
CC sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune
CC thyroiditis and insulin dependent diabetes mellitus, allergic reactions
CC and conditions such as asthma and other respiratory problems, periodontal
CC disease, osteoporosis, osteoarthritis and other tooth repair processes.
CC They may have utility in compositions used for bone, cartilage, tendon,
CC ligament and/or nerve tissue growth or regeneration as well as for wound
CC healing and tissue repair and replacement and in the treatment of burns,
CC incisions and ulcers. CD39L proteins may also be useful for proliferation
CC of neural cells and for regeneration of nerve and brain tissue, i.e. for
CC the treatment of central nervous system diseases such as Alzheimer's
CC disease, Parkinson's disease, amyotrophic lateral sclerosis, Huntington's
CC disease, peripheral nervous system diseases peripheral nerve injuries,
CC peripheral neuropathy and localised neuropathies. They are also used to
CC treat mechanical and traumatic disorders which involve degeneration,
CC death or trauma to neural cells or nerve tissue. CD39L proteins of the
CC invention are also useful to promote better or faster closure of non-
CC healing wounds, including pressure ulcers, ulcers associated with
CC vascular insufficiency and surgical and traumatic wounds. They also
CC exhibit anti-inflammatory activity and may be used to treat inflammatory
CC conditions including chronic or acute conditions), including ischaemia-
CC reperfusion injury, endotoxin lethality, arthritis, nephritis, cytokine
CC or chemokine-induced lung injury, inflammatory bowel disease or Crohn's
CC disease. The present sequence is mouse CD39L4 protein, also known as
CC nucleotide-triphosphatase (NTPase)

SQ Sequence 465 AA;

Query Match 80.7%; Score 1817.5; DB 5; Length 465;
Best Local Similarity 83.1%; Pred. No. 1.1e-166;
Matches 349; Conservative 25; Mismatches 43; Indels 3; Gaps 3;

Qy 1 MATSGTGVFFMLVSVCSAVSHRNQQTWPEGIFLSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSGAV-FMLIACVGVSTVYREQQTWEGVFLSSKCPINVSAGTFYGMFDAGSTGA 59

Qy 61 RIHYTTFVQKMPGQIPLEGVFDVSFVREGLSAFVDQPKQAGATVQGLLELVAKDSIPRSHW 120
Db 60 RIHYTTFVQKTAGQIPLEGIFDSVRPGLSAFVDQPKQAGATVQGLLELVAKDSIPRSHW 119

Qy 121 KKTVPWLKATAGLRLPHKAKALLFEVKETFRKSPFLVPKGSVIMTGDGEGIFAWTV 180
Db 120 ERTVPWLKATAGLRLPHKAKALLFEVKETFRKSPFLVPKGSVIMTGDGEGIFAWTV 179

Qy 181 NFLGQLHGRQETVGTLDLGASTQTITFLPQPEKTLQTPRGYLTSPFEMNSTYKLYTH 240
Db 180 NFLGQLHGRQETVGTLDLGASTQTITFLPQPEKTLQTPRGYLTSPFEMNSTYKLYTH 239

Qy 241 SYLFGGLKAARLATIGALETGTGHTFRSACLPRWLEAEWIFGKVYQVGNQEGVGF 300
Db 240 SYLFGGLKAARLATIGALETGTGHTFRSACLPRWLEAEWIFGKVYQVGNQEGVGF 299

Qy 301 EPCYAEVLVVRGKLHQPVEVQSGSFYAFSYYYDRAVDTMDIDYERKGGILKVEDFERKAR 360
Db 300 EPCYAEVLVVRGKLHQPVEVQSGSFYAFSYYYDRAVDTMDIDYERKGGILKVEDFERKAR 359

Qy 361 EVCNLENFTSGSPFLCMLDLSYITALLKDGFGPADSTVQLTKKVNNIETCWALGATPHL 420
Db 360 EVCNLSGSSGSPFLCMLDLYITALLKDGFGFAERHPLTAHKE-SFQHRDW-LGLGHL 417

RESULT 10
AAB53336
ID AAB53336 standard; protein; 330 AA.
XX AC AAB53336;
XX DT 09-MAR-2001 (first entry)
XX DE Human colon cancer antigen protein sequence SEQ ID NO:876.
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW identification; cytotoxic; cardioactive; neuroprotective; vulnary;
KW immunomodulatory; muscular; gynaecological; gastrointestinal;
KW nephrotropic; antiinfective; antibacterial; gene therapy; wound;
KW neural disorder; immune system disorder; muscular disorder;
KW reproductive disorder; gastrointestinal disorder; renal disorder;
KW infectious disease; cardiovascular disorder.
XX OS Homo sapiens.
XX PN WO200055351-A1.
XX PD 21-SEP-2000.
XX PF 08-MAR-2000; 2000WO-US005883.
XX PR 12-MAR-1999; 99US-0124270P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Ruben SM;
XX WPI; 2000-587534/55.
XX DR N-PSDB; AAC98093.
XX Colon cancer associated gene sequences, referred to as colon cancer
PT antigens, useful for the treatment, prevention, and diagnosis of colon
PT disorders such as colon cancer.

XX Claim 11; Page 1429-1431; 2104pp; English.

XX AAC97991 to AAC98763 encode the human colon cancer associated proteins,
CC called human colon cancer antigens, given in AAB53234 to AAB54006. The
CC human colon cancer antigens can have cytotoxic, cardioactive, muscular;
CC neuroprotective, immunomodulatory, gynaecological, gastrointestinal, and
CC vulnary, nephrotropic, antiinfective and antibacterial activities, and
CC can be used in gene therapy. The colon cancer antigen polynucleotides,
CC proteins and antibodies to the proteins are useful for the prevention,
CC treatment and diagnosis of colon disorders, such as colon cancer. The
CC polynucleotides may be used in diagnostics and research, such as for
CC chromosome identification, and as hybridisation probes. The proteins may
CC also be used to prevent diseases such as neural disorders, immune system
CC disorders, muscular disorders, reproductive disorders, gastrointestinal
CC disorders, wounds, renal disorders, infectious diseases, and
CC cardiovascular disorders. AAC98764 to AAC98772 and AAB54007 represent
CC sequences used in the exemplification of the present invention

XX SQ Sequence 330 AA;

Query Match 73.0%; Score 1645; DB 3; Length 330;
Best Local Similarity 96.3%; Pred. No. 3.3e-150;
Matches 316; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 101 AETVQGLLELVAKDSIPRSHWKTTPVYLKATAGLRLPHKAKALLFEVKETFRKSPFLVP 160
Db 3 ARAVQGLLELVAKDSIPRSHWKTTPVYLKATAGLRLPHKAKALLFEVKETFRKSPFLVP 62

Qy 161 KGSVSIWMTGDGEGIFAWTVNFTLQGLHGRQETVGTLDLGASTQTITFLPQPEKTLQET 220
Db 63 KGSVSIWMTGDGEGIFAWTVNFTLQGLHGRQETVGTLDLGASTQTITFLPQPEKTLQET 122

Qy 221 PRGYLTSPFEMNSTYKLYTHSYLFGGLKAARLATIGALETGTGHTFRSACLPRWLEAE 280
Db 123 PRGYLTSPFEMNSTYKLYTHSYLFGGLKAARLATIGALETGTGHTFRSACLPRWLEAE 182

Qy 281 WIFGVKYQVGNQEGVGFPCYAEVLVVRGKLHQPVEVQSGSFYAFSYYYDRAVDT 340
Db 183 WIFGVKYQVGNQEGVGFPCYAEVLVVRGKLHQPVEVQSGSFYAFSYYYDRAVDT 242

Qy 341 MIDYERKGGILKVEDFERKAREVCNLENFTSGSPFLCMLDLSYITALLKDGFGPADSTVQL 400
Db 243 MIDYERKGGILKVEDFERKAREVCNLENFTSGSPFLCMLDLSYITALLKDGFGPADSTVQL 302

Qy 401 LTKKVNNIETCWALGATPHLQSLGISH 428
Db 303 LTKKVNNIETCWALGATPHLQSLGISH 330

RESULT 11
AAB19881
ID AAB19881 standard; protein; 456 AA.
XX AC AAB19881;
XX DT 18-JUN-2002 (first entry)
XX DE Human CD39L2 protein.
XX Human; CD-39-like protein; CD39L2 protein; therapy; immune deficiency;
KW autoimmune disorder; multiple sclerosis; systemic lupus erythematosus;
KW rheumatoid arthritis; autoimmune thyroiditis; allergic reaction; asthma;
KW insulin dependent diabetes mellitus; periodontal disease; osteoporosis;
KW osteoarthritis; wound healing; tissue repair; Alzheimer's disease; ulcer;
KW Parkinson's disease; amyotrophic lateral sclerosis; Huntington's disease;
KW nervous system disease; nerve injury; ischaemia-reperfusion injury;
KW endotoxin lethality; arthritis; nephritis; inflammatory bowel disease;
KW Crohn's disease; virucide; antibacterial; antifungal; neuroprotective;
KW dermatological; immunosuppressive; vulnary; nontropic; anticonvulsant;
KW antiinflammatory; nephrotropic; gastrointestinal; vasotropic.
XX OS Homo sapiens.

XX US6350447-B1.
 PN 26-FEB-2002.
 XX 29-JAN-1999; 99US-00240639.
 PF 29-JAN-1999; 99US-00240639.
 PR (HYSE-) HYSEQ INC.
 XX Chadwick BP, Frischauf A;
 PI WPI; 2002-215262/27.
 DR N-PSDB; AAD31693.
 XX An isolated polypeptide with phosphohydrolase activity, designated
 PT CD39L2, useful to identify other proteins with which binding occurs or
 PT identify inhibitors and for treatment of, e.g., Alzheimer's, multiple
 PT sclerosis and osteoporosis.
 XX Claim 1; Fig 4; 101pp; English.
 PS The present invention relates to novel proteins with phosphohydrolase
 CC activity, designated CD-39-like (CD39L) proteins and polynucleotides
 CC encoding such proteins. CD39L proteins are useful to treat infectious
 CC diseases caused by viral, bacterial, fungal or other infection that may
 CC be treatable with CD39L. They are useful in the treatment of various
 CC immune deficiencies and disorders, autoimmune disorders such as multiple
 CC sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune
 CC thyroiditis and insulin dependent diabetes mellitus, allergic reactions
 CC and conditions such as asthma and other respiratory problems, periodontal
 CC disease, osteoporosis, osteoarthritis and other tooth repair processes.
 CC They may have utility in compositions used for bone, cartilage, tendon,
 CC ligament and/or nerve tissue growth or regeneration as well as for wound
 CC healing and tissue repair and replacement and in the treatment of burns,
 CC incisions and ulcers. CD39L proteins may also be useful for proliferation
 CC of neural cells and for regeneration of nerve and brain tissue, i.e. for
 CC the treatment of central nervous system diseases such as Alzheimer's
 CC disease, Parkinson's disease, amyotrophic lateral sclerosis, Huntington's
 CC disease, peripheral nervous system diseases peripheral nerve injuries,
 CC peripheral neuropathy and localised neuropathies. They are also used to
 CC treat mechanical and traumatic disorders which involve degeneration,
 CC death or trauma to neural cells or nerve tissue. CD39L proteins of the
 CC invention are also useful to promote better or faster closure of non-
 CC healing wounds including pressure ulcers, ulcers associated with
 CC vascular insufficiency and surgical and traumatic wounds. They also
 CC exhibit anti-inflammatory activity and may be used to treat inflammatory
 CC conditions including chronic or acute conditions, including ischaemia-
 CC reperfusion injury, endotoxin lethality, arthritis, nephritis, cytokine
 CC or chemokine-induced lung injury, inflammatory bowel disease or Crohn's
 CC disease. The present sequence is human CD39L2 protein
 XX Sequence 456 AA;
 SQ Query Match 44.2%; Score 996; DB 5; Length 456;
 Best Local Similarity 52.4%; Pred. No. 3.5e-87;
 Matches 204; Conservative 54; Mismatches 125; Indels 6; Gaps 4;
 QY 40 PINVSA---STLGIIMFDAGSTGRIHVYTVQKPCQLPILSEVFDSPKPGLSAFVDQ 96
 DB 62 PLGTAAAGHGVYGIIMFDAGSTGRIHVYTVQKPCQLPILSEVFDSPKPGLSAFVD 120
 QY 97 PKQAGTAVQGLLEVAKDSIPRSHWKKTPFVLXATAGLRLPEHAKALLFEVKEIRKSP 156
 DB 121 VERSAQGIREDLVAKQDIPDFWKATPLVLKATAGLRLPEHAKALLFEVKEIRKSP 180
 QY 157 FLVPKGSVSMIGQDEGLFAWTVNPLTGLQHGHRQETVGLDGGASTQITLPQPEKT 216
 DB 181 FLVGDDCVSNMGDEGVSAWITINFLTGLSKTPGGSSVGMLDGGSGTQIAFLPRVEGT 240
 QY 217 LEQTPRGVLYTSFENFNSTYKLYTHSYLGFGLKAAKLATLGALETE-GTDDGHTFRSACLPR 275

DB 241 LOASPPGVLTAIRNFNTYKLYSYLGLGMSARLAILGVGEQPAKDGKELVSPCLSP 300
 QY 276 WLEAEWIPGGVYQYGNQGEVGFPCYAEVLRVVRKGLHQPVEVQSGSFYAFSYYYDR 335
 DB 301 SFKGEWEHAEVTVRVSGOKAAASLHELCAARVSEVLQNRVHRTBEVKHVDYAFSYYYDL 360
 QY 336 AVDTDMIDYEGGILKVEDFERKAREVCDNLENFTSGSPFLQMDLSYITALLKQGFAD 395
 DB 361 AAGVGLIDAEKGGSLWGDPEIAAKYVCRITETPOSPFFSKMDLTVVSLLLQBF-GFPFR 419
 QY 396 STVLQLTKKVNNIETGVALGATPHLLQSL 424
 DB 420 SKVLKLTREKIDNVETSWALGAIFHYIDSL 448
 RESULT 12
 AAB72241
 ID AAB72241 standard; protein; 484 AA.
 XX AAB72241;
 AC AAB72241;
 DT 14-MAY-2001 (first entry)
 XX Human CD39 like protein CD39-L2 amino acid sequence.
 DE Human CD39-like protein; apyrase; NDPase; platelet function inhibitor;
 KW myocardial infarction; cerebral ischaemia; angina; arterial thrombosis;
 KW cerebral artery thrombosis; platelet aggregation; inflammation;
 KW apoptosis; autoimmune disorder; neurological disorder;
 KW Alzheimer's disease; Parkinson's disease; cancer; CD39-L2.
 OS Homo sapiens.
 XX WO200110205-A1.
 PN 15-FEB-2001.
 PD 09-AUG-2000; 2000MO-US021790.
 XX 09-AUG-1999; 99US-00370265.
 PR 11-JAN-2000; 2000US-00481238.
 PR 25-APR-2000; 2000US-00557800.
 PR 26-MAY-2000; 2000US-00583231.
 PR 30-JUN-2000; 2000US-00608285.
 XX (HYSE-) HYSEQ INC.
 XX Ford J, Mulero JJ, Yeung G;
 PI WPI; 2001-147489/15.
 DR N-PSDB; AAF63386.
 XX Polynucleotides encoding human CD39-like polypeptides, with apyrase
 PT and/or NDPase activity, which are useful in the treatment of pathological
 PT conditions caused by thrombosis (e.g. myocardial infarction) and
 PT inflammatory disorders.
 XX Claim 39; Page 162-164; 203pp; English.
 CC This invention relates to polynucleotides encoding human CD39-like
 CC polypeptides with apyrase and/or NDPase activity. The polypeptides having
 CC ATPase, including NDPase, activity are useful for inhibiting platelet
 CC function and can therefore be used in the prophylaxis or treatment of
 CC pathological conditions caused by or involving thrombosis or excessive
 CC coagulation or excessive platelet aggregation, such as myocardial
 CC infarction, cerebral ischaemia, angina, arterial thrombosis, cerebral
 CC artery thrombosis or intracardiac thrombosis, and conditions associated
 CC with venous thrombosis. CD39-L4 and CD39-L2 polypeptides are useful in
 CC modulating disease states (including platelet aggregation, inflammation
 CC and apoptosis) associated with ADP or other purinergic signalling by
 CC reducing the levels of NDPs. The polypeptides are also useful for
 CC prophylaxis or treatment of inflammation related disorders, such as
 CC disorders involving sepsis or systemic inflammatory response syndrome or

Db 427 SKVLKTRKIDNVETSWALGAIFHYIDSL 455

RESULT 14

ABJ04657

ID ABJ04657 standard; protein; 467 AA.

AC ABJ04657;

XX

DT 11-OCT-2002 (first entry)

XX

DE Protein of NOVX 15a SEQ ID No 36.

XX

XX Cytostatic; antidiabetic; anorectic; metabolic; nootropic; antilipaeamic;

KW neuroprotective; antiparkinsonian; anticonvulsant; cerebroprotective;

KW tranquilliser; neuroleptic; antidiabetic; antitumor; antiinflammatory;

KW anti-HIV; antiallergic; antirheumatic; antiarthritic; NOVX; diabetes;

KW metabolic disorder; obesity; infectious disease; Alzheimer's disease;

KW anorexia; neurodegenerative disorder; Parkinson's disorder; obesity;

KW immune disorder; haematopoietic disorder; dyslipidaemia; chronic disease;

KW metabolic syndrome X; wasting disorder; cancer; neurological disorder;

KW epilepsy; stroke; mental disorder; schizophrenic disorder; goiter;

KW vesicular transport; cystic fibrosis; gastrointestinal disorder;

KW diabetes mellitus; ulcerative colitis; AIDS; allergic reaction;

KW multiple sclerosis; rheumatoid arthritis; transgenic animal;

gene therapy.

XX

OS Unidentified.

OS

PN W0200246409-A2.

PN

XX 13-JUN-2002.

XX

PF 06-DEC-2001; 2001WO-US046586.

XX

XX 06-DEC-2000; 2000US-0251660P.

PR 12-DEC-2000; 2000US-0255029P.

PR 08-JAN-2001; 2001US-0260326P.

PR 24-JAN-2001; 2001US-0263800P.

PR 20-FEB-2001; 2001US-0269942P.

PR 24-APR-2001; 2001US-0286183P.

PR 20-AUG-2001; 2001US-0313627P.

PR 12-SEP-2001; 2001US-0318712P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Guo X, Li L, Patturajan M, Shimkets RA, Casman SJ, Malyankar UM,

PI Tchernev VT, Vernet CAM, Spytek KA, Shenoy SG, Aisbrook JP;

PI Edinger S, Peyman JA, Stone DJ, Ellerman K, Gangolli EA, Boldog FL;

PI Colman SD, Eisen AJ, Liu X, Padigaru M, Spaderna SK, Zerhusen BD;

XX

XX WPI: 2002-547774/58.

DR N-PSDB; ABT05470.

DR

XX

XX Novel isolated polypeptide, designated NOVX, useful for treating or

PT preventing cancer, diabetes, obesity, dyslipidemia, anorexia, and

PT metabolic, neurodegenerative, immune and hematopoietic disorders.

XX

XX Claim 1; Page 140; 421pp; English.

PS

XX The invention relates to an isolated polypeptide, designated NOVX,

CC comprising a sequence fully defined in the specification. The isolated

CC protein, its encoding polynucleotide or an antibody created from the

CC protein is useful in the manufacture of a medicament for treating a

CC syndrome associated with a human disease, preferably a NOVX-associated

CC disorder, or for treating or preventing a NOVX-associated disorder in a

CC subject, preferably human. The isolated protein, its encoding

CC polynucleotide or an antibody created from the protein are also useful

CC for treating or preventing metabolic disorders, diabetes, obesity,

CC infectious disease, anorexia, neurodegenerative disorder, Alzheimer's

CC disease, Parkinson's disorder, immune disorders, haematopoietic

CC disorders, and various dyslipidaemias, metabolic disturbances associated

CC with obesity, the metabolic syndrome X, wasting disorders associated with

CC chronic diseases, and cancer. The isolated protein, its encoding

CC polynucleotide or an antibody created from the protein are useful for

CC treating or preventing neurological disorders such as epilepsy, stroke,

CC mental disorders including mood, anxiety, schizophrenic disorders,

CC disorders of vesicular transport such as cystic fibrosis, diabetes

CC mellitus, goiter, gastrointestinal disorders including ulcerative

CC colitis, other conditions associated with abnormal vesicle trafficking

CC including AIDS, allergic reactions, multiple sclerosis and rheumatoid

CC arthritis. A cell comprising the vector of the invention is useful for

CC producing non-human transgenic animals. The polynucleotide of the

CC invention can be used to treat disorders by gene therapy. This sequence

CC represents one of the isolated NOVX proteins of the invention

XX

SQ Sequence 467 AA;

Query Match 44.1%; Score 993; DB 5; Length 467;

Best Local Similarity 52.2%; Pred. No. 7e-87;

Matches 203; Conservative 55; Mismatches 125; Indels 6; Gaps 4;

QY 40 PINVSA---STLYGIMFDAGSTGTHIHVYTFQVMPCQLPILEGEVFDVSKLSAFVDQ 96

DB 73 PLGTAADGHEVYGYIMFDAGSTGTRVHVQFT-RPPRETPTLTHTFKALPKLSAYADD 131

QY 97 PKQGAETVQGLLEVAKDSIPRSHMKKTPVLKATAGLRLPEHKAKALLFEVKEIFRKSP 156

DB 132 VEKSAQGIRELLDVAKQDIPDFPKATPLVLKATAGLRLPCKEAKQLLQKVEVEKASP 191

QY 157 FLVPGSVSIMTQDEGIFAWVTNPLTQGLHGRQETVGLDGGASTQITFLPQFEXT 216

DB 192 FLVGGDCVSIIMGTDEGVSAWITINFLTSLKTPGSSVGMGLDGGSTQIAFLPRVREGT 251

QY 217 LEOTPRGYLTSEFMFNSTYKLYTHSYLGFGFKAAKLATLGALETS-GLDGHTRFACLP 275

DB 252 LQASPPGYLTALRMFNRTYKLYSYLGGLMSARLAILGGVEGQPAKDKGLNSPCLSP 311

QY 276 WLEAEWIFGVKYQYGNQGEVGEPEPCYAEVLRVVRKGLHQPEVQSGSFVAFYYTDR 335

DB 312 SFKEWEHAETVYRVSGQKAAASHLCAARVSEVLQNRVHRTVEEVKGVDFYAFSYTDL 371

QY 336 AVDTMDIDYKGGILLKVEDFERKAREVCNLENFTSGSPFELCMDLSYITALLKDGFGAD 395

DB 372 RAGVGLIDABKGSLVVGDFEFAAYVCTRTLETQPSQFFSCMDLTYVSLLLQE-FGPR 430

QY 396 STVLQLTKKVNNIETGNGALGATFHLLOSL 424

DB 431 SKVLKTRKIDNVETSWALGAIFHYIDSL 459

RESULT 15

AAM93929

ID AAM93929 standard; protein; 456 AA.

XX

AC AAM93929;

XX

DT 06-NOV-2001 (first entry)

XX

DE Human polypeptide, SEQ ID NO: 4100.

XX

XX Human; full length cDNA; cDNA synthesis; oligo-capping.

XX

XX Homo sapiens.

OS

XX EP1130094-A2.

PN

XX 05-SEP-2001.

PD

XX 07-JUL-2000; 2000EP-00114089.

PF

XX 08-JUL-1999; 99JP-00194486.

PR 11-JAN-2000; 2000JP-00118774.

PR 02-MAY-2000; 2000JP-00183765.

XX

XX (HELI-) HELIX RES INST.

PA


```
Db 121 KKTPTVVKATAGLRLPEHKAALLFEVKIIFRKSPLFVLPKGSVIMDSGDEGILAWTV 180
Qy 181 NFLTGLQHGHRQETVGTLDGGASTQITFLPQPEKTLQTPRGYLTSEFNFSTYKLYTH 240
Db 181 NFLTGLQHGHRQETVGTLDGGASTQITFLPQPEKTLQTPRGYLTSEFNFSTYKLYTH 240
Qy 241 SYLGFGLKAARLATIAGLETGTGHTFRSACLPRWLEAEWIFGVKYQYGGNQBGEVGF 300
Db 241 SYLGFGLKAARLATIAGLETGTGHTFRSACLPRWLEAEWIFGVKYQYGGNQBGEVGF 300
Qy 301 EPCVAEVLVRVVGKLGHOPEEVOGRSFAFSYVYDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 301 EPCVAEVLVRVVGKLGHOPEEVOGRSFAFSYVYDRAVDTMDIDYKGGILKVEDFERKAR 360
Qy 361 EVCNLENFTSGSPFLCMLDLSYITALLKDGFGFADSTVLQ 420
Db 361 EVCNLENFTSGSPFLCMLDLSYITALLKDGFGFADSTVLQ 420
Qy 421 LQSIGISH 428
Db 421 LQSIGISH 428

RESULT 2
Q8WUB3 PRELIMINARY; PRT; 407 AA.
AC Q8WUB3 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Similar to ectonucleoside triphosphate diphosphohydrolase 5.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RX MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium.
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AC031581; BAC27461.1; -.
DR MGD; AK01321385; Entpds.
DR InterPro; IPR000407; GDAI_CD39_NTPase.
DR Pfam; PF01150; GDAI_CD39; 1.
DR KX
SQ SEQUENCE 407 AA; 45336 MW; D92A5F7DC9E9E5B CRC64;

Query Match 92.8%; Score 2089; DB 4; Length 407;
Best Local Similarity 99.2%; Pred. No. 1.8e-171; Indels 0; Gaps 0;
Matches 397; Conservative 0; Mismatches 3;

Qy 1 MATSWGTVFFMLVSVCSAVSHRNQQTWFEIGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSWGTVFFMLVSVCSAVSHRNQQTWFEIGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Qy 61 RIHYTTFVQKMPGQPLILEGEVFDSPKGLSAFVDQPKQGAETVQGLLEAKDSIPRSHW 120
Db 61 RIHYTTFVQKMPGQPLILEGEVFDSPKGLSAFVDQPKQGAETVQGLLEAKDSIPRSHW 120
Qy 121 KKTPTVVKATAGLRLPEHKAALLFEVKIIFRKSPLFVLPKGSVIMDSGDEGILAWTV 180
Db 121 KKTPTVVKATAGLRLPEHKAALLFEVKIIFRKSPLFVLPKGSVIMDSGDEGILAWTV 180
Qy 181 NFLTGLQHGHRQETVGTLDGGASTQITFLPQPEKTLQTPRGYLTSEFNFSTYKLYTH 240
Db 181 NFLTGLQHGHRQETVGTLDGGASTQITFLPQPEKTLQTPRGYLTSEFNFSTYKLYTH 240
Qy 241 SYLGFGLKAARLATIAGLETGTGHTFRSACLPRWLEAEWIFGVKYQYGGNQBGEVGF 300
Db 241 SYLGFGLKAARLATIAGLETGTGHTFRSACLPRWLEAEWIFGVKYQYGGNQBGEVGF 300
Qy 301 EPCVAEVLVRVVGKLGHOPEEVOGRSFAFSYVYDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 301 EPCVAEVLVRVVGKLGHOPEEVOGRSFAFSYVYDRAVDTMDIDYKGGILKVEDFERKAR 360

RESULT 3
Q8CD29 PRELIMINARY; PRT; 427 AA.
AC Q8CD29 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 5.
ON NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium.
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK031581; BAC27461.1; -.
DR MGD; AG01321385; Entpds.
DR InterPro; IPR000407; GDAI_CD39_NTPase.
DR Pfam; PF01150; GDAI_CD39; 1.
DR KX
SQ SEQUENCE 427 AA; 47101 MW; 6E3773C842B58477 CRC64;

Query Match 87.7%; Score 1975.5; DB 11; Length 427;
Best Local Similarity 87.4%; Pred. No. 1.2e-161; Indels 1; Gaps 1;
Matches 373; Conservative 25; Mismatches 28;
```

```
Db 301 EPCVAEVLVRVVGKLGHOPEEVOGRSFAFSYVYDRAVDTMDIDYKGGILKVEDFERKAR 360
Qy 361 EVCNLENFTSGSPFLCMLDLSYITALLKDGFGFADSTVLQ 400
Db 361 EVCNLENFTSGSPFLCMLDLSYITALLKDGFGFADSTVLQ 400

RESULT 3
Q8CD29 PRELIMINARY; PRT; 427 AA.
AC Q8CD29 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 5.
ON NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium.
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK031581; BAC27461.1; -.
DR MGD; AG01321385; Entpds.
DR InterPro; IPR000407; GDAI_CD39_NTPase.
DR Pfam; PF01150; GDAI_CD39; 1.
DR KX
SQ SEQUENCE 427 AA; 47101 MW; 6E3773C842B58477 CRC64;

Query Match 87.7%; Score 1975.5; DB 11; Length 427;
Best Local Similarity 87.4%; Pred. No. 1.2e-161; Indels 1; Gaps 1;
Matches 373; Conservative 25; Mismatches 28;
```

```
Qy 1 MATSWGTVFFMLVSVCSAVSHRNQQTWFEIGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSWGTVFFMLVSVCSAVSHRNQQTWFEIGIFLSSMCPINVSASTLYGIMFDAGSTGT 59
Qy 61 RIHYTTFVQKMPGQPLILEGEVFDSPKGLSAFVDQPKQGAETVQGLLEAKDSIPRSHW 120
Db 61 RIHYTTFVQKMPGQPLILEGEVFDSPKGLSAFVDQPKQGAETVQGLLEAKDSIPRSHW 119
Qy 121 KKTPTVVKATAGLRLPEHKAALLFEVKIIFRKSPLFVLPKGSVIMDSGDEGILAWTV 180
Db 121 KKTPTVVKATAGLRLPEHKAALLFEVKIIFRKSPLFVLPKGSVIMDSGDEGILAWTV 179
Qy 181 NFLTGLQHGHRQETVGTLDGGASTQITFLPQPEKTLQTPRGYLTSEFNFSTYKLYTH 240
Db 181 NFLTGLQHGHRQETVGTLDGGASTQITFLPQPEKTLQTPRGYLTSEFNFSTYKLYTH 239
Qy 241 SYLGFGLKAARLATIAGLETGTGHTFRSACLPRWLEAEWIFGVKYQYGGNQBGEVGF 300
Db 241 SYLGFGLKAARLATIAGLETGTGHTFRSACLPRWLEAEWIFGVKYQYGGNQBGEVGF 299
Qy 301 EPCVAEVLVRVVGKLGHOPEEVOGRSFAFSYVYDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 301 EPCVAEVLVRVVGKLGHOPEEVOGRSFAFSYVYDRAVDTMDIDYKGGILKVEDFERKAR 359
Qy 361 EVCNLENFTSGSPFLCMLDLSYITALLKDGFGFADSTVLQ 420
Db 361 EVCNLENFTSGSPFLCMLDLSYITALLKDGFGFADSTVLQ 419
Qy 421 LQSIGISH 427
Db 421 LQSIGISH 426
```

function).

DE Homo sapiens (Human).

OS Rukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Placenta;

RA Strauberg R.;

RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC025980; AAH25980.1; -

DR GO; GO:0016787; F:hydrolase activity; IEA.

DR InterPro; IPR000407; GDAI_CD39_NTPase.

DR Pfam; PF01150; GDAI_CD39_1.

KN Hydrolase.

SQ SEQUENCE 483 AA; 53119 MW; A850E5035BCDC8F CRC64;

Query Match 44.1%; Score 993; DB 4; Length 483;

Best Local Similarity 52.2%; Pred. No. 7.7e-77;

Matches 203; Conservative 55; Mismatches 125; Indels 6; Gaps 4

Qy 40 PINVSA---STLYGIMFDAGSTGTHVYTPVQKMPGQLPILGEVFDVSKGLSAFVDQ 96

Db 89 PLGTAADGHEVYGMFDAGSTGTRVHVQFT-PPRETPTLTTHETFKALPKGLSAYADD 147

Qy 97 PKQGAETVQGLLEAVKDSIPRSHWKTTPVLKATAGLRLPEHAKALLFFVXBIFRKSP 156

Db 148 VEKSAQGIREDLDAKQIDPPDFWKATPLVLKATAGLRLPCKEAKQLQKXKVEVFKASP 207

Qy 157 FIVPKGSVIMTQDREGIPAWVTVPFLGQLHGHQETVGTLDLGGASTQTTFIPQFPT 216

Db 208 FLVGGDVCVIMNGTDEGVSAWITINPLTGLSLKTPGSSVGMGLDGGGSGTQIAFLPRVEGT 267

Qy 217 LEQTPRGYLTSEMFNSTYKLYTHSYGLGFLKAARLATLGALETB-GTDTGHTFSACLPR 275

Db 268 LQASPSYLTALRMFNTYKLYSYGLGLMSARLATLGGVEGPAKDGKELVSPCLSP 327

Qy 276 WLAEAWIPGGVKYQYGGNQEVEGPEPCYARVLRVVRGKLHOPVEVQSGSVAFSYYYDR 335

Db 328 SFKGEWEHAETVYRSGQKAAASLHELCAARVSEVLQNRVHRTVEVGVDFVAFSYYYDL 387

Qy 336 AVDTOMIDYKGGILKVEDPERKAREVCNDLENFTSGSPFLCMDLSYITALLKDGFGFAD 395

Db 388 AGVGLIDAEKGGSLVVGDFEIAAKYVCTRLETPQSPFSSOMDLTYVSLLLQ8-FGPPR 446

Qy 396 STVLOLTKVNNIETGALGATFFHLQSL 424

Db 447 SKVLKTRKIDNVETSWALGAIFHYIDSL 475

RESULT 6

Q725B5

ID Q725B5 PRELIMINARY; PRT; 484 AA.

AC Q725B5;

DT 01-OCT-2003 (TREMBlrel. 25, Created)

DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE CD39L2 nucleotidase.

GN ENTDPD6.

OS Homo sapiens (Human).

OC Rukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RA Ivanenkov V.V., Murphy-Piedmonte D.M., Kirley T.L.;

RT "bacterial Expression, Characterization, and Disulfide Bond

RT Determination of Soluble Human NTPDase6 (CD39L2) Nucleotidase:

RT Implications for Structure and Function.";

RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY327581; AAP92131.1; -

SQ SEQUENCE 484 AA; 53274 MW; D00A5D2915DF36CE CRC64;

Query Match 87.5%; Score 1969.5; DB 11; Length 427;

Best Local Similarity 87.1%; Pred. No. 3.9e-161;

Matches 372; Conservative 24; Mismatches 30; Indels 1; Gaps 1;

Qy 1 MATSWGTVPFMLVVSVCVSAVSHRNQOTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT 60

Db 1 MATSWGAV-FMLIACVGSVTVPYREQOTWFEGLFLSSMCPINVSAGTYPYGMFDAGSTGT 59

Qy 61 RIHVYTFVQKMPQLPILSGEVPDSVKPGLSAPVDPQKQGAETVQGLLEAVKDSIPRSHW 120

Db 60 RIHVYTFVQKMPQLPILSGEVPDSVKPGLSAPVDPQKQGAETVQGLLEAVKDSIPRSHW 119

Qy 121 KKTTPVVKATAGLRLPEHAKALLPEVKEIIFRKSFPFLVPGKSVIMTQDREGIPAWVTY 180

Db 120 ERTPPXKATAGLRLPEHAKALLPEVKEIIFRKSFPFLVPGKSVIMTQDREGIPAWVTY 179

Qy 181 NFLTGLHGHQETVGTLDLGGASTQITFLPQFETLEQTPRGYLTSEMFNSTYKLYTH 240

Db 180 NFLTGLHGHQETVGTLDLGGASTQITFLPQFETLEQTPRGYLTSEMFNSTYKLYTH 239

Qy 241 SYLFGGLKARLATLGALETGDTGHTFSAKLPWLEAWETFGGVKYQYGGNQEVEGVP 300

Db 240 SYLFGGLKARLATLGALETGDTGHTFSAKLPWLEAWETFGGVKYQYGGNQEVEGVP 299

Qy 301 EPCYAEVLRVVRGKLHOPVEVQSGSVAFSYYYDRAVDTMDIDYKGGILKVEDPERKAR 360

Db 300 EPCYAEVLRVVRGKLHOPVEVQSGSVAFSYYYDRAVDTMDIDYKGGILKVEDPERKAR 359

Qy 361 EVCNDLENFTSGSPFLCMDLSYITALLKDGFGFADSTVQLTKVNNIETGALGATPHL 420

Db 360 EVCNDLENFTSGSPFLCMDLSYITALLKDGFGFADSTVQLTKVNNIETGALGATPHL 419

Qy 421 LQSLGTS 427

Db 420 LQSLGTS 426

RESULT 5

Q8TAS7

ID Q8TAS7 PRELIMINARY; PRT; 483 AA.

AC Q8TAS7;

DT 01-JUN-2002 (TREMBlrel. 21, Created)

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 24, Last annotation update)

DE Ectonucleoside triphosphate diphosphohydrolase 6(Putative

Qy 421 LQSLGTS 427

Db 420 LQSLGTS 426

RESULT 5

Q8TAS7

ID Q8TAS7 PRELIMINARY; PRT; 483 AA.

AC Q8TAS7;

DT 01-JUN-2002 (TREMBlrel. 21, Created)

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 24, Last annotation update)

DE Ectonucleoside triphosphate diphosphohydrolase 6(Putative

Qy 421 LQSLGTS 427

Db 420 LQSLGTS 426

RESULT 5

Q8TAS7

ID Q8TAS7 PRELIMINARY; PRT; 483 AA.

AC Q8TAS7;

DT 01-JUN-2002 (TREMBlrel. 21, Created)

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 24, Last annotation update)

DE Ectonucleoside triphosphate diphosphohydrolase 6(Putative

Qy 421 LQSLGTS 427

Db 420 LQSLGTS 426

RESULT 5

Q8TAS7

ID Q8TAS7 PRELIMINARY; PRT; 483 AA.

AC Q8TAS7;

DT 01-JUN-2002 (TREMBlrel. 21, Created)

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 24, Last annotation update)

DE Ectonucleoside triphosphate diphosphohydrolase 6(Putative

Qy 421 LQSLGTS 427

Db 420 LQSLGTS 426

RESULT 5

Q8TAS7

ID Q8TAS7 PRELIMINARY; PRT; 483 AA.

AC Q8TAS7;

DT 01-JUN-2002 (TREMBlrel. 21, Created)

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 24, Last annotation update)

DE Ectonucleoside triphosphate diphosphohydrolase 6(Putative

Qy 421 LQSLGTS 427

Db 420 LQSLGTS 426

RESULT 5

Q8TAS7

ID Q8TAS7 PRELIMINARY; PRT; 483 AA.

AC Q8TAS7;

DT 01-JUN-2002 (TREMBlrel. 21, Created)

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 24, Last annotation update)

DE Ectonucleoside triphosphate diphosphohydrolase 6(Putative

Qy 421 LQSLGTS 427

Db 420 LQSLGTS 426

RESULT 5

Q8TAS7

ID Q8TAS7 PRELIMINARY; PRT; 483 AA.

AC Q8TAS7;

DT 01-JUN-2002 (TREMBlrel. 21, Created)

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 24, Last annotation update)

DE Ectonucleoside triphosphate diphosphohydrolase 6(Putative

Qy 421 LQSLGTS 427

Db 420 LQSLGTS 426

RESULT 5

Q8TAS7

ID Q8TAS7 PRELIMINARY; PRT; 483 AA.

AC Q8TAS7;

DT 01-JUN-2002 (TREMBlrel. 21, Created)

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 24, Last annotation update)

DE Ectonucleoside triphosphate diphosphohydrolase 6(Putative

Qy 421 LQSLGTS 427

Db 420 LQSLGTS 426

RESULT 5

Q8TAS7

ID Q8TAS7 PRELIMINARY; PRT; 483 AA.

AC Q8TAS7;

DT 01-JUN-2002 (TREMBlrel. 21, Created)

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 24, Last annotation update)

DE Ectonucleoside triphosphate diphosphohydrolase 6(Putative

Qy 4

```
Query Match 44.1%; Score 993; DB 4; Length 484;
Best Local Similarity 52.2%; Pred. No. 7.8e-77;
Matches 203; Conservative 55; Mismatches 125; Indels 6; Gaps 4;

QY 40 PINVSA---STLYGIMFDAGSGTGRHIVYTFVQKMPGQLPILGEVDSVKPGLSAFVDQ 96
DB 90 PLGTAADGHEVYFYGIMFDAGSGTGRVHVQFT-RPPRETPTLTHTFKALPKGLSAYADD 148
QY 97 PQQAEITVQGLLEVAKDSIPRSHMKKTTPVLKATAGLRLPEHAKALLFEVKIKRSP 156
DB 149 VEKSAQIGRELLDVAKQDIPDFWKATPLVLKATAGLRLPEHAKALLQKVKVFRASP 208
QY 157 FLVPKGSVSIWQDEGIPAMVTNNFLGQLHGHQETVGLDGGASTQITFLPQPKT 216
DB 209 FLVGGDDCVSIMNGTDEGVSAMITINFLTGLSKTPGSSVGMGLDGGSTQIAFLPRVEGT 268
QY 217 LEQTPRGYLTSPFEMNSTYKLYTHSYLGFGLKAARLATIAGALETE-GTGDGHTFASACLP 275
DB 269 LQASPPGYLTALRMENRTYKLYSYLGLGLMSARLAILGGVQOPAKGKELVSPCLSP 328
QY 276 WLEAEWIFGGVKYQYGGNQEVEGPEPCVAEVLVVRGKLHQPBEVQSGSYAFSYYYDR 335
DB 329 SPKGEWEHAEVTVRVSGQKAAASLHELCAARVSEVLQNRVHRTBEVKHVDYAFSYYYDL 388
QY 336 AVDTMDIDYKGGILKVEDFERKAREVCNLENFTSGSPFLCMLDSYITALLKDGFGFAD 395
DB 389 AAGVGLIDAEKGGSLVGDPEFAIAKVVCRTLETQPOSSPFCMDLTAVSLLLQGE-FGFPR 447
QY 396 STVQLTKKNNIETGALGATFHLLQSL 424
DB 448 SKVLKLRKIDNVETSWALGAIFHYIDSL 476

RESULT 7
Q8N3H3 PRELIMINARY; PRT; 503 AA.
AC Q8N3H3
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein (Fragment).
GN DKF2761919.5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Amysdala;
RA Anserge W., Wirkner U., Mewes H.W., Weil B., Wiemann S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL834158; CAD38864.1; -
DR InterPro; IPR000407; GDA1_CD39_NTPase.
DR Pfam; PF01150; GDA1_CD39_1.
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 503 AA; 54763 MW; 1639333F9139DBD9F CRC64;

Query Match 44.1%; Score 993; DB 4; Length 503;
Best Local Similarity 52.2%; Pred. No. 8.2e-77;
Matches 203; Conservative 55; Mismatches 125; Indels 6; Gaps 4;

QY 40 PINVSA---STLYGIMFDAGSGTGRHIVYTFVQKMPGQLPILGEVDSVKPGLSAFVDQ 96
DB 109 PLGTAADGHEVYFYGIMFDAGSGTGRVHVQFT-RPPRETPTLTHTFKALPKGLSAYADD 167
QY 97 PQQAEITVQGLLEVAKDSIPRSHMKKTTPVLKATAGLRLPEHAKALLFEVKIKRSP 156
DB 168 VEKSAQIGRELLDVAKQDIPDFWKATPLVLKATAGLRLPEHAKALLQKVKVFRASP 227
QY 157 FLVPKGSVSIWQDEGIPAMVTNNFLGQLHGHQETVGLDGGASTQITFLPQPKT 216
DB 157 FLVPKGSVSIWQDEGIPAMVTNNFLGQLHGHQETVGLDGGASTQITFLPQPKT 216

RESULT 8
O76268 PRELIMINARY; PRT; 461 AA.
AC O76268
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NTPase protein (LD11641P).
GN NTPASE OR CG3059.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foaier C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush E., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Khamel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lin Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Moberg C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M.C., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasmadge T.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage D., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
```

"The genome sequence of Drosophila melanogaster.";
Science 287:2185-2195(2000).

```
RT  The genome sequence of Drosophila melanogaster.";  
RL  Science 287:2185-2195(2000).  
RN  (1)  
RP  SEQUENCE FROM N.A.  
RX  MEDLINE=98341119; PubMed=9676430;  
RA  Chadwick B.F., Frischauf A.-M.;  
RT  "The CD39-like gene family: identification of three new human members  
RT  (CD39L2, CD39L3, and CD39L4), their murine homologues, and a member of  
RT  the gene family from Drosophila melanogaster.";  
RL  Genomics 50:357-367(1998).  
RN  (3)  
RP  SEQUENCE FROM N.A.  
RC  STRAIN=Berkley;  
RA  Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
RA  Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
RA  Gonzalez M., Guarini H., Li P., Liao G., Miranda A., Mungall C.J.,  
RA  Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,  
RA  Yu C., Lewis S.E., Rubin G.M., Celnikier S.;  
RL  Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR  EMBL; AF003581; AAC1182.1; -;  
DR  EMBL; AF041048; AAC39133.1; -;  
DR  EMBL; AY061134; AAL28682.1; -;  
DR  FlyBase; FBgn0024947; NTPase.  
DR  InterPro; IPR000407; GDAL_CD39_NTPase.  
DR  Pfam; PF01150; GDAL_CD39; 1.  
SQ  SEQUENCE 461 AA; 50845 MW; 27D00321F91A9DD1 CRC64;  
  
Query Match 30.8%; Score 693.5; DB 5; Length 461;  
Best Local Similarity 39.1%; Pred. No. 4.9e-51;  
Matches 156; Conservative 73; Mismatches 123; Indels 47; Gaps 12;  
  
QY 49 YGIMFDAGSTGTRIHVTVFVKMPGQLPILEGEVDSVKPGLSAPVDQPKQGAETVOGILL 108  
DB 79 YAAIDAGSTGSRVLAYKFNRSFIDNKLVLVYELFKERAPGLSSPADNPAGAHSIKULL 138  
  
QY 109 EVAKDSIPRSHWKKTPVVLKATAGLRLPEHKAKALLFEVKIPFKSPFLPKGVSIMT 168  
DB 139 DEARAFIPKEHSSTPLVKATAGLRLPASKAENILNAVRLFAKSEPSVDMDAVEIMD 198  
  
QY 169 GODEGIFAWVTNFLTQGLHGRQETVGTLDGGASTQITF-----LPQFKTLEQTP 221  
DB 199 GDEGIFSWFTNFLTGLRLSKTNQ--AAALDGGGSTQVTFSTPTDQDPVYDKYMHV- 255  
  
QY 222 RGYLTSPFMFNSTVKLYTHSYLGFGLKAARLATLGALETEG--TDGHTFRSACL-PRWLE 278  
DB 256 ---VTSSKKIN---VFTHSYLGLGLMARHAFV-----THGYKKEPTVLESVCVNPILAN 304  
  
QY 279 AEWIFGGVKYQYGGNQE-----VGFEPCVAEVLVRVHGKL-----HQPEEVQVQSGF 326  
DB 305 RTWTYGNVQYKVSCKENGKSSAEQPIVDFDAC----LELVKSKVMPLVKPKSPFTLKQHAV 360  
  
QY 327 YAFSYTYDRAVDMDIDYKGGILKVEDFERKAREVC--DNLENFTSGSPFLCMDLSYIT 384  
DB 361 AAFSYTYFERAIESGLVDPLAGETTVEATRKQAQELCAIPNDE-----QPFNCFDLTFTIS 415  
  
QY 385 ALLKQGFADSTVLQLTKKVNNIEFWGALGATEHLLQS 423  
DB 416 TLIRGEGLNDGKKIKLYKKIDGHEISWALGCAVNLTS 454
```

RESULT 9

```
Q9VQI8  
ID Q9VQI8 PRELIMINARY; PRT; 464 AA.  
AC  
CD  
DT 01-MAY-2000 (T-EMBLrel. 13, Created)  
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (T-EMBLrel. 17, Last annotation update)  
DE NTPASE protein.  
GN NTPASE OR C30359.  
OS Drosophila melanogaster (fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.
```

```
OX NCBI_TaxID=7227;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX STRAIN=Berkley;  
RA MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Nikols G.L.G.,  
RA April J.P., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burris J.M., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry K.C., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,  
RA Fostler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J.J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasmann D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
DR EMBL; AE003581; AAF51181.1; -;  
DR FlyBase; FBgn0024947; NTPase.  
DR InterPro; IPR000407; GDAL_CD39_NTPase.  
DR Pfam; PF01150; GDAL_CD39; 1.  
SQ SEQUENCE 464 AA; 51119 MW; 71D057ABE5AB613D CRC64;
```

Query Match 30.8%; Score 693.5; DB 5; Length 464;
Best Local Similarity 39.1%; Pred. No. 4.9e-51;
Matches 156; Conservative 73; Mismatches 123; Indels 47; Gaps 12;

```
QY 49 YGIMFDAGSTGTRIHVTVFVKMPGQLPILEGEVDSVKPGLSAPVDQPKQGAETVOGILL 108  
DB 82 YAAIDAGSTGSRVLAYKFNRSFIDNKLVLVYELFKERAPGLSSPADNPAGAHSIKULL 141  
  
QY 109 EVAKDSIPRSHWKKTPVVLKATAGLRLPEHKAKALLFEVKIPFKSPFLPKGVSIMT 168  
DB 142 DEARAFIPKEHSSTPLVKATAGLRLPASKAENILNAVRLFAKSEPSVDMDAVEIMD 201  
  
QY 169 GODEGIFAWVTNFLTQGLHGRQETVGTLDGGASTQITF-----LPQFKTLEQTP 221  
DB 202 GDEGIFSWFTNFLTGLRLSKTNQ--AAALDGGGSTQVTFSTPTDQDPVYDKYMHV- 258  
  
QY 222 RGYLTSPFMFNSTVKLYTHSYLGFGLKAARLATLGALETEG--TDGHTFRSACL-PRWLE 278  
DB 259 ---VTSSKKIN---VFTHSYLGLGLMARHAFV-----THGYKKEPTVLESVCVNPILAN 307  
  
QY 279 AEWIFGGVKYQYGGNQE-----VGFEPCVAEVLVRVHGKL-----HQPEEVQVQSGF 326  
DB 308 RTWTYGNVQYKVSCKENGKSSAEQPIVDFDAC----LELVKSKVMPLVKPKSPFTLKQHAV 363
```


Search completed: July 1, 2004, 13:48:40
Job time : 39.5 secs

```

RESULT 15
Q8TGH6 PRELIMINARY; PRT; 599 AA.
AC Q8TGH6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Guanosine diphosphatase (EC 3.6.1.42).
GN GDAL.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
NCBI_TaxID=5476;
RX [1]
RN SEQUENCE FROM N.A.
RP Herrero A.B.;
RT "The Golgi Gpase of the fungal pathogen Candida albicans Affects
RT morphogenesis, Glycosylation and cell wall properties.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AJ421721; CAD18870.1; -.
DR GO; GO:0004382; F:guanosine diphosphatase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000407; GDAL_CD39_NTPase.
DR Pfam; PF01150; GDAL_CD39_1.
DR PROSITE; PSIC1238; GDAL_CD39_NTPase; 1.
DR Hydrolase.
SQ SEQUENCE 599 AA; 65952 MW; D5EC3A4F6E1B646A CRC64;

Query Match 22.8%; Score 513.5; DB 3; Length 599;
Best Local Similarity 32.3%; Pred No. 2.4e-35;
Matches 142; Conservative 67; Mismatches 147; Indels 83; Gaps 15

QY 49 YGIMFDAGSTGTRIHHVTVFKMKGQPLPILGEVFDSVKGLGAFVDQPKQGAETVGGLL 108
DB 163 YVVMIDAGSTGSRVHVVFNTCV--KPPQLLSSEFENLKPGCLSFDTITVGAAKSLDPL 220
QY 109 EVAKDSIPRSHWKKTPVVLKATAGLRLLPEHKALLPEVKETPRKS-PP-LVPEKGSVSI 166
DB 221 EVALKVPKQKSCIPVAVKATAGLRLLGETSKAILDDEVRSLEKDYPPAVVSEGGISI 280
QY 167 NTGQDEGIPIAWTVNPLTGLQHLG-HRQETVGTGLDGGASTQITPLPQFEKLTQTPRGYL 225

```


OM protein - protein search, using sw model

Run on: July 1, 2004, 13:43:42 ; Search time 16.5 Seconds

(without alignments)
2495.150 Million cell updates/sec

Title: US-10-091-085-7

Perfect score: 2252

Sequence: 1 MATSWGVTFVFLVMSVCVCSA.....ETGVALGATFLLQLSLGISH 428

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_78.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	616.5	27.4	479	2 T23508	hypothetical prote
2	535.5	23.8	556	2 T39109	probable guanosine
3	488.5	21.7	518	2 A40732	guanosine-diphosph
4	484	21.5	455	2 S48859	nucleoside triphos
5	441.5	19.6	454	2 JC4616	apyrase (EC 3.6.1.
6	426.5	18.9	516	2 G84442	probable nucleosid
7	398	17.7	483	2 D86276	hypothetical prote
8	382.5	17.0	485	2 T34147	hypothetical prote
9	374.5	16.6	557	2 T16696	hypothetical prote
10	371	16.5	630	2 S50463	hypothetical prote
11	366	16.3	510	2 I56242	lymphoid cell acti
12	345	15.3	572	2 T40956	probable nucleotid
13	332.5	14.8	405	2 E86276	hypothetical prote
14	279	12.4	1052	2 T04329	hypothetical prote
15	242	10.7	508	2 C86276	7A19.33 protein -
16	148	6.6	628	2 A55421	nucleoside-triphos
17	108.5	4.8	1019	2 T40913	probable cell divi
18	106.5	4.7	535	2 F97910	glucan 1,6-alpha-g
19	105.5	4.7	553	1 G1BPSV	gene 1 protein - s
20	104.5	4.6	535	2 C95040	glucan 1,6-alpha-g
21	100.5	4.5	1151	2 H71347	hypothetical prote
22	99.5	4.4	774	2 T14555	DNA polymerase hom
23	99	4.4	711	2 S66749	hypothetical prote
24	99	4.4	726	2 AB0122	probable ferric si
25	98.5	4.4	4427	2 PN0637	polyketide synthas
26	98	4.4	497	1 JC2192	subtilisin-like pr
27	96.5	4.3	287	2 T40138	probable ribosomal
28	96	4.3	477	2 J50597	t-plasminogen acti
29	95.5	4.2	301	2 E71482	phosphatidylserine

hypothetical prote
H4-transporting tw
hypothetical prote
probable serine/th
probable type II D
coproporphyrinogen
acetylornithine am
vril protein - Dic
hypothetical prote
patched protein -
hypothetical prote
cytochrome o ubiqu
capular polysacch
betaine aldehyde d
protein F35H12.4
glycine hydroxymet

ALIGNMENTS

RESULT 1

T23508
hypothetical protein K08H10.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C:Accession: T23508
R:Gardner, A.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19750
A:Accession: T23508
A:Status: Preliminary; translated from GB/EMBL/DBDJB
A:Molecule type: DNA
A:Residues: 1-479 <MIL>
A:Cross-references: EMBL:Z83113; PIDN: CAB05544.1; GSPDB:GN000023; CESP:K08H10.4
A:Experimental source: clone K08H10
C:Genetics:
A:Gene: CESP:K08H10.4
A:Map position: 5
A:Introns: 36/3; 83/3; 189/1; 300/2; 412/3
C:Superfamily: nucleoside triphosphatase chromatin-associated

Query Match	27.4%;	Score	616.5;	DB 2;	Length	479;			
Best Local Similarity	34.1%;	Pred. No.	7.3e-44;						
Matches	149;	Conservative	82;	Mismatches	167;	Indels	39;	Gaps	16;
QY	9	FFMLVVS--CVCSAVSHRNQQTWFGIFLSSMCPINVSASTLYGIMFDAGSTGTRIHVYT	66						
DB	6	PSILLISFSPSLLSVVTKQY-WCHGDGVN----NQHTCRFTTIVIDAGSTGTRLHLKY	60						
QY	67	FVQK-----MPGQILPILEGVFDSVKPGLSAFVDQPKQCAETVQGLLEVAKDSIPRSH	119						
DB	61	FIHDPALASHGNPFK---VEKEIFQEVKPKGLSSPAKSPSSAADSLEPLLQRAKQVPHFM	117						
QY	120	WKKTVPVLKATAGLRLPEHKAKALLFEVKE-IPRKSPLVPKGSVSIIMTQDQGIQPAWV	178						
DB	118	WEXTPTILKATAGLRLPGDMADDILESVEERIFNSGFAPFPDAVNVNFGSDGVQSWF	177						
QY	179	TYNFLTQQLH-----GHR---QETVGTLDLGGASTOITFLPQPKLTLPQTRGYLTSTF	229						
DB	178	TNLTLELTFTDEPTVGHKPAARHSAAPDLGGSTQLTYPNNEAVPSEHV-GYERDID	236						
QY	230	MFNSTVTKLYTHSYLGFGLKAARLITAGLETG--TDGHTFRSACLPRNLE-AEMIFGV	286						
DB	237	FFGHIRLFTHSFLGNGLIAARLNIL-QLETDNEIESTHQLITSCMEGYQLTEWEY-AL	294						
QY	287	KYQYGNQGEVGFPCPYAEVLVRVR-GKLHQPEEVQVGSFYAPSVYVYDRAVDTMIDYE	345						
DB	295	KP-WNINGSSSHSFSCYGTTRKNFVSESIIMHLRELKGSFVFLPSFYFFDRALNSGLVKGN	353						
QY	346	KGGLKVEDFERKAREVC-----DNLENFTSGSPFLCMDLSYITALLKDGPGFADSTVLQL	401						
DB	354	EGGKIELRPFKEAABEACREKTEIDDGSHNMPWQCLDITYIVSLRDGCVQFEDNQPLVL	413						

QY 402 TKKNNIETGHALGATP 418
DB 414 AKKINGMEVSGQGLAF 430

RESULT 2
T39109
Probable guanosine-diphosphatase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T39109
R:Barrell, B.G.; Rajandream, M.A.; Quail, M.; Seegar, K.; Harris, D.
submitted to the EMBL Data Library, October 1999
A:Reference number: Z21828
A:Accession: T39109
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-556 <BAR>
A:Cross-references: EMBL:AL121741; PIDN:CAB57338.1; GSPDB:GN000056; SPDB:SPAC824.08
A:Experimental source: strain 972h-; cosmid c824
C:Genetics:
A:Gene: SPDB:SPAC824.08
A:Map position: 1

Query Match 23.8%; Score 535.5; DB 2; Length 556;
Best Local Similarity 34.1%; Pred. No. 5e-37;
Matches 143; Conservative 61; Mismatches 158; Indels 57; Gaps 13;

QY 49 YGIMFDAGSTGTRTHVTEVQKMPGQLPILGEVDSVKGLGSAFVDPQKQGAETVQGL 108
DB 134 YVLMIDAGSTGSRVHVTFQPNCFNS--PKLEEFKMFIEFGLSGFADPGRAASLDPLL 191

QY 109 EVAKDSIPRSHWKTTPVVKATAGLRLLPEHKAKALLPEVKEIFRKS-PFLVPRKGSVIM 167
DB 192 DYAMENVEEVRRCSPIAVATAGLRLLTGESEAKILKSVRQHLNDYPPFIVKDGVSIL 251

QY 168 TGQDEGIFANVTNFLTQGLHGH-ROETVGTDLGGASTQITLPOFEKTELEQTPRG--- 223
DB 252 EGSMEGIYAMTINLTGLTGKATHSTVAVMDLGGASTQLVPEPRASGESLVMDGDKH 311

QY 224 YLTSFEMFNSTYKLYTHSYLGFGLKAARL-----ATLGALETEGTGHTFRSAC 272
DB 312 YVLIDYN--GEQYELVQSHLGYGLKEARKLIHKFVLNNAEALKESLELLG-DSTSIIHPC 368

QY 273 L-----PWLEAEWTFEGVKYQYGGNQBEGVFBCYAEVLVR-----GKLRQP 318
DB 369 LHLNASLTHPDSKSEAEVVFVGPESLAHLSLQCRGIAEKALYKDKNCPVRPCSFNGVHQP 428

QY 319 ---EEVQKGSFYAPSYVYDRAVDTMDIYEKGGILKVEDFERKAREVC----- 363
DB 429 KFTETFTDSPIYLSIYDR-----MISLGNPSTFTIEDMKYLANSVCSGPTYWQDAFSL 483

QY 364 -DNLENTSGSPFLCMLDSYITALLKDGFGFADSTVLQTKKNNIETGHTFRSAC 421
DB 484 TDALKELKE-EPWCLDLNTYMISLLSVGYEIPNRRQLHTAKKIDNKELGWLGLASLSML 541

RESULT 3
A40732
Guanosine-diphosphatase (EC 3.6.1.42) - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YEL042W
C:Species: Saccharomyces cerevisiae
C>Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 02-Sep-2000
C:Accession: A40732; B40732; S50502
R:Abeion, C.; Yanagisawa, K.; Mandon, E.C.; Hausler, A.; Moremen, K.; Hirschberg, C.B.;
J. Cell Biol. 122, 307-323, 1993
A:Title: Guanosine diphosphatase is required for protein and sphingolipid glycosylation
A:Reference number: A40732; MUID:93308137; PMID:8391537
A:Molecule type: DNA
A:Accession: A40732
A:Residues: 1-518 <ABE>
A:Cross-references: EMBL:L19560; NID:g349392; PIDN:AAA34656.1; PID:g349393

A:Note: sequence extracted from NCBI backbone (NCBIN:134708, NCBIP:134711)
A:Accession: B40732
A:Molecule type: protein
A:Residues: 125-144;238-257;276-281;366-374;399-412 <AB2>
R:Mulligan, J.T.; Dietrich, F.S.; Hennessey, K.M.; Sehl, P.; Komp, C.; Wei, Y.; Tayl
submitted to the EMBL Data Library, February 1993
A:Reference number: S30812
A:Accession: S30837
A:Molecule type: DNA
A:Residues: 1-518 <MUL>
A:Cross-references: GB:U18779; EMBL:L10830; NID:5603625; PIDN:AAB65000.1; PID:560363
R:Dietrich, F.S.
A:Description: The sequence of S. cerevisiae cosmids 8199, 8334, and 9871.
A:Reference number: S50491
A:Accession: S50502
A:Molecule type: DNA
A:Residues: 1-518 <DIE>
A:Cross-references: EMBL:U18779; NID:5603625; PIDN:AAB65000.1; PID:5603637; MIPS:YEL
C:Genetics:
A:Gene: SGD:GDAL
A:Cross-references: SGD:S0000768; MIPS:YEL042W
A:Map position: 5L
A:Function:
C:Description: hydrolase
C:Superfamily: nucleoside triphosphatase chromatin-associated
C:Keywords: glycoprotein; Golgi apparatus; hydrolase; transmembrane protein
F:10-24/Domain: transmembrane #status predicted <TMM>
F:41,280,335/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 21.7%; Score 488.5; DB 2; Length 518;
Best Local Similarity 33.0%; Pred. No. 5e-33;
Matches 143; Conservative 56; Mismatches 159; Indels 75; Gaps 15;

QY 49 YGIMFDAGSTGTRTHVTEVQKMPGQLPILGEVDSVKGLGSAFVDPQKQGAETVQGL 107
DB 93 YVLMIDAGSTGSRVHVTFQPNCFNS--PTLLDERFDMLEPGLSGSFDTDSVGAANSLDPL 149

QY 108 LEVAKDSIPRSHWKTTPVVKATAGLRLLPEHKAKALLPEVKEIFRKS-PFLVPRKGSVIM 165
DB 150 LKVAKYVPIKARSCPTVAVKATAGLRLLGDAKSKILSAVRDHLKEDYPPFVVEGDGVS 209

QY 166 IMTQDEGIFANVTNFLTQGL--HGHRQETVGTDLGGASTQITLPOFEKTELEQTPRG 223
DB 210 IMGDEEGVFANITNYLLGNIGANGPKLPTAAVFDLGGSGTQIVPEFPFPIKERNVDE 269

QY 224 YLTSFEMFNSTYKLYTHSYLGFGLKAAR-----LATLGALETEGTGHTFRSAC 272
DB 270 HKFDLKFGEENTYLYQPSHLGYGLKEGRKNVSVLVENALKDGLKILKGNNTKTHQLSSPC 329

QY 273 LPRWLEA--EWI-----FGSVKYQYCG-----NORGEVGFEPFCYAEVL 309
DB 330 LSPKVNATNWKVILSKETTYTIDFIDPDEPSGAQCRFLTDEILNKDAQCQSPCSF---- 385

QY 310 VVRGKLHQPEEV-----QRGSFYAFSYDRAVDTDM-IDYKGGILKVEDFERKAREVC- 363
DB 386 ---NGVHQPISLVRTPKESNDIYFVYDTRPLGMLPSF-----TLNELNDLARIUCK 436

QY 364 -----DNLENTSGSPFLCMLDSYITALLKDGFGFADSTVLQTKKNNIET 410
DB 437 GSETNWSVPSGAGSLDELSDSHF-CLDLSQVSLHTGTGDIPIQREURTKRKANKRI 495

QY 411 GWALGATFHLLQS 423
DB 496 GNCGLGASLPLAKA 508

RESULT 4
S48859
nucleoside triphosphatase precursor, chromatin-associated - garden pea
C:Species: Pisum sativum (garden pea)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 29-Sep-1999
C:Accession: S65147; S48859

QY 402 TKKNNIETGHALGATP 418
DB 414 AKKINGMEVSGQGLAF 430

RESULT 2
T39109
Probable guanosine-diphosphatase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T39109
R:Barrell, B.G.; Rajandream, M.A.; Quail, M.; Seegar, K.; Harris, D.
submitted to the EMBL Data Library, October 1999
A:Reference number: Z21828
A:Accession: T39109
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-556 <BAR>
A:Cross-references: EMBL:AL121741; PIDN:CAB57338.1; GSPDB:GN000056; SPDB:SPAC824.08
A:Experimental source: strain 972h-; cosmid c824
C:Genetics:
A:Gene: SPDB:SPAC824.08
A:Map position: 1

Query Match 23.8%; Score 535.5; DB 2; Length 556;
Best Local Similarity 34.1%; Pred. No. 5e-37;
Matches 143; Conservative 61; Mismatches 158; Indels 57; Gaps 13;

QY 49 YGIMFDAGSTGTRTHVTEVQKMPGQLPILGEVDSVKGLGSAFVDPQKQGAETVQGL 108
DB 134 YVLMIDAGSTGSRVHVTFQPNCFNS--PKLEEFKMFIEFGLSGFADPGRAASLDPLL 191

QY 109 EVAKDSIPRSHWKTTPVVKATAGLRLLPEHKAKALLPEVKEIFRKS-PFLVPRKGSVIM 167
DB 192 DYAMENVEEVRRCSPIAVATAGLRLLTGESEAKILKSVRQHLNDYPPFIVKDGVSIL 251

QY 168 TGQDEGIFANVTNFLTQGLHGH-ROETVGTDLGGASTQITLPOFEKTELEQTPRG--- 223
DB 252 EGSMEGIYAMTINLTGLTGKATHSTVAVMDLGGASTQLVPEPRASGESLVMDGDKH 311

QY 224 YLTSFEMFNSTYKLYTHSYLGFGLKAARL-----ATLGALETEGTGHTFRSAC 272
DB 312 YVLIDYN--GEQYELVQSHLGYGLKEARKLIHKFVLNNAEALKESLELLG-DSTSIIHPC 368

QY 273 L-----PWLEAEWTFEGVKYQYGGNQBEGVFBCYAEVLVR-----GKLRQP 318
DB 369 LHLNASLTHPDSKSEAEVVFVGPESLAHLSLQCRGIAEKALYKDKNCPVRPCSFNGVHQP 428

QY 319 ---EEVQKGSFYAPSYVYDRAVDTMDIYEKGGILKVEDFERKAREVC----- 363
DB 429 KFTETFTDSPIYLSIYDR-----MISLGNPSTFTIEDMKYLANSVCSGPTYWQDAFSL 483

QY 364 -DNLENTSGSPFLCMLDSYITALLKDGFGFADSTVLQTKKNNIETGHTFRSAC 421
DB 484 TDALKELKE-EPWCLDLNTYMISLLSVGYEIPNRRQLHTAKKIDNKELGWLGLASLSML 541

RESULT 3
A40732
Guanosine-diphosphatase (EC 3.6.1.42) - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YEL042W
C:Species: Saccharomyces cerevisiae
C>Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 02-Sep-2000
C:Accession: A40732; B40732; S50502
R:Abeion, C.; Yanagisawa, K.; Mandon, E.C.; Hausler, A.; Moremen, K.; Hirschberg, C.B.;
J. Cell Biol. 122, 307-323, 1993
A:Title: Guanosine diphosphatase is required for protein and sphingolipid glycosylation
A:Reference number: A40732; MUID:93308137; PMID:8391537
A:Molecule type: DNA
A:Accession: A40732
A:Residues: 1-518 <ABE>
A:Cross-references: EMBL:L19560; NID:g349392; PIDN:AAA34656.1; PID:g349393

R:Shieh, H.L.; Tong, C.G.; Thomas, C.; Roux, S.J.

Plant Mol. Biol. 30, 135-147, 1996
A:Title: Light-modulated abundance of an mRNA encoding a calmodulin-regulated, chromatin
A:Reference number: S65141; MUID:96197404; PMID:8616230
A:Accession: S65147
A:Molecule type: mRNA
A:Residues: 1-455 <HS2>
A:Cross-references: EMBL:Z32743; NID:G563611; PIDN:CA83655.1; PID:G563612
A:Superfamily: nucleoside triphosphatase chromatin-associated
C:Keywords: nucleus

Query Match 21.5%; Score 484; DB 2; Length 455;

Best Local Similarity 33.4%; Pred. No. 9.8e-33;

Matches 139; Conservative 62; Mismatches 161; Indels 54; Gaps 14;

QY 49 YGIMFDAGTGTIRIHVTFVQKMPGQPLILEG-EVFDVSKPGLSAFVDPKQGAETVQGL 107

DB 44 YAVVFDAGTGTIRIHVTFVQKMPGQPLILEG-EVFDVSKPGLSAFVDPKQGAETVQGL 102

QY 108 LEVAKDSIPRSHWKKTPVVLKATAGRLLLPEHAKALLFEVKEIF-RKSPFLVPKGSVSI 166

DB 103 LEQAEVDPDLPKTPVRLGATAGLLNGDASEKILQSVRDMLSNRSTFVQPDVSI 162

QY 167 MTQDQSGIFAWVTNVLFTQLHGHROETVGTLDLGASTQITFLPOFEKTLBOTPR----- 222

DB 163 IDGTQSGYLWTVNVALGNLKKYTKTVGVIDLGGSVQMAVAS-KXTAKNAPKRVADG 221

QY 223 --GYLTSFEMFNSTYKLYTHSYLGRGLKAAALATLGALETGEGTGHFTPS--ACLPRLWE 278

DB 222 DPEYIKVVLKGPIDLYVHSYLFHFGREASRAEILKL-----TPRSPNPLLGNFN 272

QY 279 AEWIFGVKYQYGGNGEGEVGFPCYAEVLVVRGKHLHPEEVQVQSGP----- 326

DB 273 GIYVSGEFGKATAYTSG-ANFNKCKNTIRKAL--KLNVPCPYQNCFTFGIWNNGGNGNQ 329

QY 327 ---YAEYVYDRAVDMDIDYKGG-ILKVEDEPKAREVCD-NLENFTSGSPFL----- 376

DB 330 KMLFASSTFFLPEDTGMWDASTPFIILRPDIETKAKACALNPEDEKASTTFPDLKXNV 389

QY 377 ----CMDSYITALLXGFGFADSTVLQTKKVN----IETGVALGATFHLLOSL 424

DB 390 ASYVCMDLIVQVLLVDFGLDPLQKITSGKEIEYDAIVEAAWPLGNVAEISAL 445

RESULT 5

JC4616

aprase [BC 3.6.1.5] precursor - potato

N:Alternate names: adenylylprophosphatase; ATP-diphosphohydrolase

C:Species: Solanum tuberosum (potato)

C:Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 28-May-1999

C:Accession: JC4616; PC4147

R:Handa, M.; Guidotti, G.

Biochem. Biophys. Res. Commun. 218, 916-923, 1996

A:Title: Purification and cloning of a soluble ATP-diphosphohydrolase (Apyrase) from pot

A:Reference number: JC4616; MUID:96158985; PMID:8579614

A:Accession: JC4616

A:Molecule type: mRNA

A:Residues: 1-454 <HAN>

A:Cross-references: GB:U58597; NID:gl381632; PIDN:AA802720.1; PID:gl381633

A:Accession: PC4147

A:Molecule type: protein

A:Residues: 59-95/96-131/132-160/236-253/332-345 <HA2>

A:Experimental source: tubers

A>Note: The authors translated the codon GCA for residue 215 as Gly

C:Comment: This enzyme belongs to a family of E-type Apyrases, and it catalyzes the hydro

lyzed in starch synthesis.

C:Genetics:

A:Gene: rrop1

C:Superfamily: nucleoside triphosphatase chromatin-associated

C:Keywords: glycoprotein; hydrolase; transmembrane protein

F:1-30/Domain: signal sequence #status predicted <sig>

F:8-25/Domain: transmembrane #status predicted <tm>

F:31-454/Product: ATP-diphosphohydrolase #status predicted <MAT>
F:44-65/Region: actin-heat shock protein 70-hexokinase beta-phosphate binding
F:192-212/Region: nucleotide binding #status predicted
F:192-212/Region: actin-heat shock protein 70-hexokinase gamma-phosphate binding
F:390-410, 427-446/Region: nucleotide binding #status predicted
F:151, 262/Binding site: carboxylic (Asn) (covalent) #status predicted

Query Match 19.6%; Score 441.5; DB 2; Length 454;

Best Local Similarity 28.6%; Pred. No. 3.7e-29;

Matches 130; Conservative 78; Mismatches 178; Indels 69; Gaps 16;

QY 25 NOOTWFEIGIFLS-----SMCPINVSA-----STLYGIMFDAGSTGTIRIHVY 65

DB 3 NONSHFIILAILFLVPLSLSKVNAQIPLRRLSHSEHYAVIFDAGSTGSRVVP 62

QY 66 TPVQKPGQPLILEG-EVFDVSKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHWKTTP 124

DB 63 RFDEKL-GLLPIGNINIEYFMATEPGLSSYAEADPKAAANSLRPLLDGAEGVVPQELQSTP 121

QY 125 VVLKATAGRLLLPEHAKALLFEVKEIFR-KSPFLVPKGSVSIIMTQDQEGIFAWVTNVL 183

DB 122 LELGATAGRLMLKGAASKILOAVNLVKNQSTHSHKQDQWTLIDGTQEGSYMAAINYL 181

QY 184 TQQLHGHROSTVGTLDLGASTQITFL---POFEKTLR-OTPRGYLTSFEMFNSTYKLYT 239

DB 182 LGNLGKDYKSTTATIDLGGSVQMAVATSNQFAKAPQNEDEGEPYVQCKHLSKDYNLV 241

QY 240 HSYLFGFLKAAALATLGALETGEGTGHFTPSACLPRLWEAWIFGVKQYQYGGNGEGEVG 299

DB 242 HSYLNYGQLAGRAEIFKASRNES-----NPCALEGCGGYSGYGVGYKVPKPKGS-- 292

QY 300 FEPYAEVLVVRG--KLHQPVEVQVQSGF-----YAFSYTYDRAVDTDMI 342

DB 293 ---SNKRCERLTERALKINAKNIEECFTFNGVWGGGDKGNHASSFFYDGAQVGIV 349

QY 343 DYE-KGGLKVEDFERKAREVCD-NLENFTS-----GSPFLCMDLSYITALLKQDGF 392

DB 350 DTKFPSALAKPIQLYNAKAVACQTNVADIKSIFFKTDQNPITYLCMDLIYTYLLVDGFG 409

QY 393 F---ADSTVLQTKKVN-IETGVALGATFHLLOS 423

DB 410 LNPHEKITVIEDVQYKYLVGAAWPLGCAIDLVS 444

RESULT 6

G84442

probable nucleoside triphosphatase [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 18-Nov-2002

C:Accession: G84442

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Unayam, L.; Tallon

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Vente

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: G84442

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-516 <STO>

A:Cross-references: GB:AB002093; NID:g3461821; PIDN:AAC32915.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g02970

A:Map position: 2

C:Superfamily: nucleoside triphosphatase chromatin-associated

Query Match 18.9%; Score 426.5; DB 2; Length 516;

Best Local Similarity 30.4%; Pred. No. 8.2e-28;

Matches 126; Conservative 63; Mismatches 153; Indels 73; Gaps 13;

QY 42 NVSASTLYGIMFDAGSTGTIRIHVTFVQKMPGQPLILE-GEVFDVSK--PGLSAFVDPQ 97

```
Db 61 NLRGSLRYSVWIDGSGTCTRIHVFGY--RIESGKPVFEFRGANYASIKLHPGLSADFADDP 118
Qy 98 KQGAETVOGLLEVAKDSDIPRSHWAKTTPVWLKATAGLRLPEHRAKALLFEVKIFRKSPP 157
Db 119 DGASVSLTELVEFAKVRPKGMWLETVEVLWATAGMLLELPVQSKILGVARRVLKSSGP 178
Qy 158 LVPKGSVINTGQDEGIFAMVTVNFTGLQHLGHRQETVGTLDIGGASTQTTPFL-----P 211
Db 179 LFRDEWASVIGSGDEGVYAWVANFALGSLGGDFLKTGTGIVELGASAAQVTFVSSEMPMP 238
Qy 212 QFEKTLQTPRGYLTSPFEMNSTYKLYTHSYLGLGKAAALATLGAAL-----ETRGTD 264
Db 239 EFSRTI-----SFG--NVTNLYSHSLFHFGQNAADHKLWGLSLSRDHSNAVEPTR 287
Qy 265 GHTFRSACLPR-----WLEAEWTFGGVQYQYGNQGEVGFPCYAEVURVV 311
Db 288 EKIFTPCAPKGYNDANTQKLSGLLAESRLSDSPQAGCN-----YSQCRSAALTIL 341
Qy 312 RGKLHQPEEVQGSFYAFSYVYDRAVDTDMI-----DYEGGILKVEDPERKAREV 362
Db 342 QDGNRLIITAGTSFIFFGLEGKAWLSMISASERFGEDWSK---LRVADPSLHREDL 398
Qy 363 CDNLNFTSGSPFLCMLSYTALLKDGFGF-ADSTVLQTLTKKYNNTISTGAWLGA 416
Db 399 LR-----YCFSSAYIVSLHDLTGIPLDDEIRIGYANQAGDIPLDLWALGA 442

RESULT 7
D86276
hypothetical protein F7A19.34 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 18-Nov-2002
C:Accession: D86276
R:Theologos, A.; Becker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
Nansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
C:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
Xer, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D86276
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-483 <STO>
A:Cross-references: GB:AE005172; NID:95080801; PIDN:AAD39311.1; GSPDB:GN00141
A:Map position: 1
C:Genetics:
C:Superfamily: nucleoside triphosphatase chromatin-associated

Query Match 17.7%; Score 398; DB 2; Length 483;
Best Local Similarity 27.4%; Pred. No. 1.9e-25;
Matches 119; Conservative 73; Mismatches 161; Indels 82; Gaps 17;

Qy 32 GFLSSMCPINVASLYGIMFDAGSTGTRIHV--YTFVQKAPQQLPILGEVFD-----84
Db 59 GSLLSRCKLA-----YSVLIDAGSSGTRVHVFGESEKPE-----VDFGEKH 103
Qy 85 ----SVKPGLSAFVDPQKQGAETVOGLLEVAKDSIPRSHWAKTTPVWLKATAGLRLPEHK 140
Db 104 YANLKLTPGLSSYADNPEGASVSTKLVEPAKQIPKRMFRSDIRLMATAGMELLEVPV 163
Qy 141 AKALLFVKEIFRKSPPFLVPRKGSVINTGQDEGIFAMVTVNFTGLQHLGHRQETVGTLDL 200
Db 164 QEQLLEVTRVLRSSGFREMANVIGSGDEGIYSMTANYALGSLGTDPLETGVIEL 223
Qy 201 CGASTQTTLFPQFQKTLQTPRGYLTSPFEMNSTYKLYTHSYLGLGKAAALATLGALET 260
Db 224 CGASAAQVTFV-----SSRHVPFYSRTIAYGNIYTSYSHSFLDYGDAAALKKLEKLN 278
```

```
Qy 261 EG---TDGHTFRSACLPRWLEAEWTF--GGVKYQVQ-----GNQGEVGFPCYAE 306
Db 279 SANSTVDG--VVEDPCTPK---GYIVDTNSKNYSSGFLADESKLKGSLQAAGNFSKCRSA 333
Qy 307 VLRVVR-----GKLHQPEEVQGSFYAFSYVYDRAVDTDMI DYEGGILKLV- 352
Db 334 TPALLKGEKENCLEHYECSIGSTFTPD--LQGSFLATASFYTAKPEFL--BKGMISELI 389
Qy 353 -----EDFERKAREVCDNLNFTSGSPFLCMLSYTALLKDGFGFPA--DSTVLQUT 402
Db 390 PAKRYCGEWSKILLEYPTTDEEVLRG---YCFSAAYTISMLHDSGLIALDDESITYAS 446
Qy 403 KK-VNNIETGAWLGA 416
Db 447 KAGEKHIPLDMLGA 461

RESULT 8
T34147
hypothetical protein C33H5.14 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 17-Mar-2000
C:Accession: T34147
R:Bradshaw, H.; Stellyes, L.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid C33H5.
A:Reference number: Z21482
A:Accession: T34147
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-483 <BRA>
A:Cross-references: EMBL:U41007; PIDN:AAA82272.1; CESP:C33H5.14
A:Experimental source: strain Bristol N2
A:Genetics:
A:Gene: CESP:C33H5.14
A:Introns: 22/1; 83/1; 120/2; 167/1; 269/3; 399/3
C:Superfamily: nucleoside triphosphatase chromatin-associated

Query Match 17.0%; Score 382.5; DB 2; Length 485;
Best Local Similarity 27.5%; Pred. No. 3.8e-24;
Matches 119; Conservative 84; Mismatches 153; Indels 77; Gaps 18;

Qy 49 YGIMFDAGSTGTRIHVYTFVQKAPQQLPILGEVFD-----VKPGLSAFVDPQKQAE 102
Db 25 YGVICDAGSGSTRLFVYT-LKPLSGGLTNIDTLIHSESEFVVKVKTPLGSLSGDKPQVVE 83
Qy 103 TVQGLLEVAKDSIPRSHWAKTTPVWLKATAGLRLPEHRAKALLFEVKEIFRKSPPFL-VPK 161
Db 84 YLTPLRFAEHLIPYQLGETDLIPATAGMLLPEAKDAIKMLONGLKSVTALRVSD 143
Qy 162 GSVSINTGQDQBGIFAMVTVNFTGLQHLGHRQETVGTLDIGGASTQTTPLPQPEK-----215
Db 144 SNRIIDGAWEGIYSIAVNYILGRPKENDSKVGMIDMGASVQIAFIAEXESYNGG 203
Qy 216 TLEQTPRGYLTSPFEMNSTYKLYTHSYLGLGKAAALATLGALETGDTGHTFRSACLPR 275
Db 204 NVYEINLGSITNEDTK--YKIYSTTFLGYGANEGLKKYENSLVKSNS-----NDSCSPR 257
Qy 276 WLEAEWIFGQVKYQYGNQGEVGFPCYAEVLRVVRGKLHOPE-----319
Db 258 GLNR--LIG-----EFTVNGTB--WDVCLAQVSSLI-GDKAQPSCNPCTFLNRVIAPSV 308
Qy 320 EVORGSGFYAFS--YYDRAVDTDMI DYEGGILKVEDPERKAREVC-----DNLENFTSGSP 374
Db 309 NLSTVOLYGFSEYWTTS-----NFGSGGEVHYQKPTDEVRYKQCKMNDIQDGFKRNE 362
Qy 375 F-----LCMDLSYTALLKDGFGFADST--VLQTLTKKYNNTISTGAWLGA-----416
Db 363 FPNADIERLGTNCFKAAWTSLVLDGFN-VDXTKHLFQSVLKIAGEEMQWALGAMLYHSK 421
Qy 417 --TFHLLQSLGIS 427
Db 422 DLKENLLEQLEVA 434
```

RESULT 9

T16696
 Hypothetical protein R07E4.4 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: T16696
 R:Miller, N.
 submitted to the EMBL Data Library, October 1995
 A:Description: The sequence of *C. elegans* cosmid R07E4.
 A:Reference number: Z18561
 A:Accession: T16696
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-557 <MIL>
 A:Cross-references: EMBL:U39652; NID:g1049390; PID:g1049394; PIDN:AAA0403.1; CESP:R07E4
 A:Experimental source: strain Bristol N2
 C:Genetics:
 A:Gene: CESP:R07E4.4
 A:Introns: 39/1; 67/1; 145/2; 192/1; 244/3; 317/3; 451/3; 486/2

Query Match 16.6%; Score 374.5; DB 2; Length 557;
 Best Local Similarity 28.3%; Pred. No. 2.2e-23;
 Matches 126; Conservative 70; Mismatches 159; Indels 91; Gaps 21;
 QY 49 YGIMPDAGSTGTHHYTFVQKMPGQQLPILGEVFDSD-----VKQCLSAFYVDOPKQAE 102
 DB 44 YGVICDAGSTGTLFVNWISTSDSELIQIEPVYDNKPMKKISPLGLSTFGTKPAQAAE 103
 QY 103 TVQGLLEAVKDSIPRSHWKTTPVVLKATAGLRLLPEH-----KAKALLPEVKEIFRKSPP 157
 DB 104 YLRPLMELABRHPEEKRPVTFVIFATAGWRLLPDEYVLIGQKEAVL---KNLRNKLPK 160
 QY 158 L-----VPGKSVSMTGQDEGIFAWVTNFLTQGLH-----GH-RQTVGTLDL 200
 DB 161 ITSMQVLKHIRIEGKWEIYSWIAVYALGKFNKTATLDPFGTSPAHARQKTGVGMDM 220
 QY 201 GGASTQITF-LPOFEK-----TLQTPRGYLTSEFENSTYKLVTHSVLGF---GLKAAR 251
 DB 221 GGASQAIAFLPDTSDSSINVENINLGCRDSDLFK--YKLFVTFYLGYGVEGIRKYE 278
 QY 252 LATIGALTEGTGHTFRSACLPRWLEAEWIFGKVKYQYGN--QSG----- 296
 DB 279 HMLSLKLDQ--NGTVIQDDCMPLNLHKT-----VILENGENFVRGTGNWNTCSNEVKK 331
 QY 297 ----EVCPECYAEVLVRVWGLKHOPE-EVQSGFYAFSYVDRAVDTMDIDYKESGLK 351
 DB 332 LLNPSSSEVCKAEAKCYFGAVPAPSPILSNIEMYGFSEYWTSTEDV-----LGLGQYD 387
 QY 352 VEDFERKAREVCD-----NLENFTSGP-----FLCMDSLYITALLKDGFGFADST 397
 DB 388 AENIAKKTQYCSKRWSTIQAESKKQLYPRADSERLTQCFSKAWITSVLHDGFS-VDKT 446
 QY 398 --VLQUTKKNNTIETGHALCA-TFHL 420
 DB 447 HNKFSQSVSTIAGQEVQWALGAMTYHM 472

RESULT 10

S50463
 Hypothetical protein YER005w - yeast (*Saccharomyces cerevisiae*)
 C:Species: *Saccharomyces cerevisiae*
 C>Date: 28-May-1993 #sequence_revision 24-Feb-1995 #text_change 19-Apr-2002
 C:Accession: S50463
 R:Dietrich, F.S.
 submitted to the EMBL Data Library, December 1994
 A:Description: The sequence of *S. cerevisiae* cosmids 9537, 9581, 9495, 9867, and lambda
 A:Reference number: S50433
 A:Accession: S50463
 A:Molecule type: DNA
 A:Residues: 1-630 <DIE>
 A:Cross-references: EMBL:U18778; NID:g603592; PID:g603597; GSPDB:GNC00005; MIPS:YER005W

C:Genetics:
 A:Gene: SGD:YND1; MIPS:YER005W
 A:Cross-references: SGD:S0000807
 A:Map position: 5R

Query Match 16.5%; Score 371; DB 2; Length 630;
 Best Local Similarity 26.0%; Pred. No. 5.2e-23;
 Matches 116; Conservative 75; Mismatches 164; Indels 92; Gaps 17;
 QY 49 YGIMPDAGSTGTHHYTF-----VQMPGQPLILEGE-----VFDSDVQGLSA 92
 DB 10 FGVIVDAGSGSRHVFVKWQDTESLHATNQDSQILQSVPHIQEKDWT-F-KLMPGLSS 68
 QY 93 FVDQPKQGAET-VQGLLEAVKDSIPRSHWKTTPVVLKATAGLRLLPEHKAALLFEV-KE 150
 DB 69 FEKKFQDAYKSHIKELLDFAKNIPESHWSGCPVFIQIAGWRLLPQDIQSSILDLGCG 128
 QY 151 IPRKSPFLVP--KGSVSMITGQDEGIFAWVTNFLTQGLHGRQE-----TVGTLDLGA 203
 DB 129 LKHPAEFLVEDCSAQIQVIDGSETGLYGLWGLNLYLGHFNDPNPSVDHFTFGFMDMGA 188
 QY 204 STQITFLPQPEKTL-----QTPRGYLTSEFENSTYKLVTHSYLQGLKAAEL 252
 DB 189 STQIAFAPDGSGEIARHRDDIATIFLRSVNGDLQKWDVFSV-----WLGFGANQAR 241
 QY 253 ATLGAL-----ETEGTDGHTFRSACLPRWLEAEWIFGKVKYQYGNQGEVGPPE 301
 DB 242 RYLAQLINTLPENTYNDYENDPSTENLNDPCMPGSGSTDFEFKDTIFHIAGSGNYEQCTK 301
 QY 302 PCYAEVLR-----VVRGKLHOPE-EVQSGFYAFSYVDRAVDTMDIDYKESGLK 352
 DB 302 STYPLLLKNMPDDEPCLFNGVHAPRIDFANDKFIGTSEYWTANDV-----FKLGGEYNF 357
 QY 353 EDEFEKAREVCN-----LENFTSG-----SPFL---CMDSLYITALLKDGFGA--- 394
 DB 358 DRESKSLREFCNSNTQILANSKGYNSIPENFLKDACFKGNWLVNILEGFDMPRIDV 417
 QY 395 -----DSTVLQUTKKNNTIETGHALG 415
 DB 418 DAENVNDRPLFQSVKEKVERELSWTLG 444

RESULT 11

I56242
 lymphoid cell activation antigen - human
 C:Species: *Homo sapiens* (man)
 C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Nov-2002
 C:Accession: I56242
 R:Maliszewski, C.R.; Delespesse, G.J.; Schoenborn, M.A.; Armitage, R.J.; Fanslow, M.
 J. Immunol. 153, 3574-3583, 1994
 A:Title: The CD39 lymphoid cell activation antigen. Molecular cloning and structural
 A:Reference number: I56242; MUID:95015846; PMID:7930580
 A:Accession: I56242
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-510 <RES>
 A:Cross-references: GB:I573813; NID:g765255; PIDN:AA82152.1; PID:g765256
 C:Superfamily: nucleoside triphosphatase chromatin-associated

Query Match 16.3%; Score 366; DB 2; Length 510;
 Best Local Similarity 28.2%; Pred. No. 1e-22;
 Matches 125; Conservative 72; Mismatches 172; Indels 74; Gaps 19;
 QY 32 GTFPLSMCPIVTSASTLYGIMFDAGSTGTHHYTFVQKMPGQPLILEGEVFDSDVSK-PGL 90
 DB 36 GLTONKALPENVK---YGVILDAGSSHTSLYIKWPAEKENDTGVVHVQVBCRVKPGGI 91
 QY 91 SAFPVQPKQGAETVQGLLEAVKDSIPRSHWKTTPVVLKATAGLRLL---PEHKAKALLFE 147
 DB 92 SKFVQKVNIEIGLYLTDCHERAREVIFRSQHSQSTPVLGATAGWRLLMESEELADRVLDV 151
 QY 148 VKEIFRKSPLVPKGSVSLMTGDSGIFAWVTNFLTQGL-----HGHQRTVV 195

Db 152 VERLSNVPF--DFOGARIITQBEHGYGWITNYLLGKFSQKTRWFSIVPVETNNQETTF 209
QY 196 GTLDLGGASTITFLPOFEXLTQTPRGYLTSEFMFNSTYKLYTHSYLGFLGKAARLATL 255
Db 210 GALLDGGASTQVTFVPO--NQTB--SPDNL--QFRLYKGVYVYTHSYFLCVKQDQALWQKL 266
QY 256 GALTETGDTGHTRESACL-----PRLAEWIFGGVYKVGCGQGE 297
Db 267 -AKDIQVANEILLDPCHFPGYKVVVNSLYTKTPCKTFEMILPQOFEIQIGN----- 321
QY 298 VGFEPCYAEVLRV-----VGKLLHQPEEVQVGSFYAFSYYYDRAVDTMIDYE 345
Db 322 --YQCHQSILELNTSYCPYSCAFENGIFLPPLQGDGFAFSAP--YFVWKFLN---LTSE 375
QY 346 KGGILKYVEDPERK--AREVCNLENFTSG--SPFL--CMDLSVITALLKDGFGF--ADS-T 397
Db 376 KVSEKTEEMKFKCAPWEIKTSYAGVKEKYLSEYCSGTVILSLILLOGHFHTADSWE 435
QY 398 YLQTKKNNIETGWAIGATFHL 420
Db 436 HIHFPIQSGDAGWTLYMLNL 458

RESULT 12
T40856
Probable nucleotide phosphatase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T40856
R:Ramsperger, U.; Pohl, T.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, October 1999
A:Reference number: 221952
A:Accession: T40856
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-572 <RAM>
A:Cross-references: EMBL:AL121783; PIDN:CAB57847.1; GSPDB:GN00068; SPDB:SPCC11E10.05c
A:Experimental source: strain 972h-; cosmid c11E10
C:Genetics:
A:Gene: SPDB:SPCC11E10.05c
A:Map position: 3

Query Match 15.3%; Score 345; DB 2; Length 572;
Best Local Similarity 25.3%; Pred. No. 7e-21;
Matches 114; Conservative 69; Mismatches 156; Indels 112; Gaps 16;
QY 49 YGIMFDAGSTGTRTHVYTF-----VQKMPGOLPILGEVFD-----SVKPGLSAFVD 95
Db 5 YGIFIDAGSGSRLLIISWDYDSDSLSDKVKKLPLEITGIGDGRWLSKVQPGISSFAN 64
QY 96 QPKQ--GAETVQGLLEVAKDSIPRSHMKTPVVLKATAGURLL--PEHKAKALLFEYKELPR 153
Db 65 NPKVVGKGLKELLDFAAHAIKPDVHKETPVFLSATAGMELLGVDAQNKILSHACKYIKK 124
QY 154 KSPFLVP--KGSVSIMTQDEGIFAWTVNFLTGLQHGHRQETVGTLDLGGASTQITFLP 211
Db 125 NYDFDIINCNSIRVIDGKAEQMGVGLATYLLKTLERKDTSTVGFLDMGASVQLAP-- 182
QY 212 QFEKTLQTPRGVYLTSEPMFNST-----YKLTHSYLGFLGKAARLATLGA-- 258
Db 183 -----ELPPSOLKATYKDSISTVHIGLQNGQGLYPLFVTTWLGFCANEAYRVILGLI 235
QY 259 ETE--GTGHTFRSACLPRLEASWIFGVYKVGQGG-----NOGEVGFPCY 304
Db 236 ESENGKVGNTLSDPCSIR--GRYDIDGIEPAGTGDLKQCLKLYNLNKKDPCSNDFCN 293
QY 305 AEVLVRVVGKLLHQPEEVQVGSFYAFSYYYDRAVDTMDIYKGGILKLVEDFERKAREVC- 363
Db 294 PDGISI-----PPVDPAFTEFVGVSEFTYTTNDV-----FDMGSGVHFNFYKVKVDEYCG 343
QY 364 -----DNLENFTSGPFLKQWLSVITALLKDGFGFADSTV---- 398
Db 344 TEWETMLSRILYNKLTSTDENKLEK-----LCFKASWALNVLHGEFDVPSKNTSSND 396

QY 399 -----LQTKKNNIETGWAIG 415
Db 397 AKDGLSVIPAYHSPFTSLEKIERTEVSWTLG 427

RESULT 13
E86276

hypothetical protein F1417.1 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
C:Accession: E86276
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Ali
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, Z.A.; Iaros, J.S.; Maiti, R.; Marzi
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tall
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: E86276
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-405 <STO>
A:Cross-references: GB:AB005172; NID:G7262666; PIDN:AAF43924.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 14.8%; Score 332.5; DB 2; Length 405;
Best Local Similarity 26.3%; Pred. No. 4.8e-20;
Matches 110; Conservative 64; Mismatches 157; Indels 87; Gaps 15;
QY 48 LYGIMPAGSTGTRIHVYTFVQKPGOLPILGEVFDSPKVLGSFAFVQPKQGAETVQGL 107
Db 3 VFGVWPESGXP-----VDFGEBHYASL-----KLSPLGSSVADNPEGASVSVTKL 48
QY 108 LEVAKDSIPRSHMKTPVVLKATAGURLLPEHKAKALLFEYKELPRKSPFLVPGKSVSIM 167
Db 49 VEPAGRIPIKGLKXSDIRLMATAGNLLDVPVQEQILDVTRVLESSGFKQDEWATVI 108
QY 168 TQDDEGIFAWTVNFLTGLQHGHRQETVGTLDLGGASTQITFLPQPEKTLQTPRGVYLS 227
Db 109 SGTDEGIYAVVAVNHALGSLGSDPLKTTGIVELGGASAOVTFVPS-----EHVPPBFSRT 163
QY 228 PERFNSTYKLYTHSYLGFLGKAARLATLGALETE---GTGHTFRSACLPR----- 275
Db 164 ISYGNVSTIYSHSFLDPGQDABDKLLSLQNSVAASTGDGIVEDPCTPKGYIYTHSQ 223
QY 276 -----WLEAEWIF--GGVYKVGQNGQGEVGFPCYAEVLVRVVR--GKLLHQPEEVQVGSFYA 328
Db 224 KDSGFLSBSKFKASLQVQAGD-----FTKRSATLAMLQEGKEN----- 265
QY 329 PSYYVDRAVDTMDIYKGGILKLVEDF-----ERKAREVCMDL 366
Db 266 -CAVHCISGISTTPNIGSFLATENFFHTSPFGKEKEWSEMLAGKRCFGEWSKL 324
QY 367 -ENF--TSGSPFL---CMGLSYITALLKDGFGPA--DSTVILQTKV--NNIETGWAIGA 416
Db 325 KEKYPTTKYLRHYCFSSAYIISMLHDSLGVALDDERIKYASKACKENIPLDWAIGA 382

RESULT 14
T04439

hypothetical protein T10816.150 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 30-Apr-1999
C:Accession: T04439
R:Bevan, M.; Benes, V.; Rechmann, S.; Borkova, D.; Ansoorge, W.; Bancroft, I.; Mewes,
submitted to the Protein Sequence Database, April 1998
A:Reference number: Z15359

Query Match 10.7%; Score 242; DB 2; Length 508;
Best Local Similarity 23.2%; Pred. No. 2.8e-12;
Matches 109; Conservative 78; Mismatches 161; Indels 122; Gaps 21;
QY 12 LVVSCVCSAV-----SHRNOQTWFGIPUSSMCPINVSASTYIGIMFDAGSTGT 60
DB 35 VIVACVTIAGLGLLFIGYSILSRNR-----VSLHYSVIIDGSSSGT 77
QY 61 RIHYVTFVQKPGQLPIL-GEVDSVK--PGLSAFVDQPKQGAETWGLLEVAKDSIP 116
DB 78 RVHVFY--RIESGRPVDFGENTASLSPGLSAYADNPGVSESUTELVEFAKRVH 135
QY 117 RSHWKTTPVVLKATAGLRLPEHKAALLFEVKEIFRKSPPPLVPKGSVIMTQDEGIFA 176
DB 136 KGLKAKSDIRLMATAGMLLPLVQEQILDVTRVLRSSGF-----DPRDE---- 181
QY 177 WYVNFPLGQLHGHQETVGTLDG-----CASTQITFLPQFEXTLQTPRGYLTSPM 230
DB 182 WASV---ISEILENFQDLAKVYMLGLLIMRSVRLEVFV-----STELVPSEFRTLAY 233
QY 231 FNSTYKLYTHSYLGFGLKAARLATLGA---ETECTDGHTRFSAACLPRLWLEMI---- 282
DB 234 GNVSYNLISHFLDFGQDAQOBKLSLSSLYNSAANSTGSGIVDPDPCIPKGYILETNLQKDL 293
QY 283 --PGVKYQYQGNQGEVGFPCYAEVLVRV-----GKLPQEEVORGSTY 327
DB 294 PGFLADKGTATLQAGNFSECSAFAFMLEQEKGCYKRCSTGSIPTFN--LQGSFL 351
QY 328 AFSYVY-----DRAVDTDMI-----DYEKGILKVDPERKAREVCNLENF 369
DB 352 ATENEFHTSKFPGLGEKWLSEMILAGKPCGEHNSK---LKVYPTPKD-----ENLLRY 404
QY 370 TSGSPFLCMDSYITALLKDGFGFA-DSTVLQITKKV--NNIETGWLGA 416
DB 405 -----CPSSAYIISMLHSLDGLVALDDERIKYASKAGEEDIPLDWALGA 447

Search completed: July 1, 2004, 13:49:26
Job time : 17.5 secs

A;Accession: T04439
A;Molecule type: DNA
A;Residues: 1-1052 <BEV>
A;Cross-references: EMBL:AL021687
A;Experimental source: cultivar Columbia; BAC clone T18B16
C;Genetics:
A;Map position: 4
A;Introns: 79/3; 118/3; 217/1; 295/3; 396/3; 531/2; 815/3
A;Note: T18B16-150
Query Match 12.4%; Score 279; DB 2; Length 1052;
Best Local Similarity 25.1%; Pred. No. 6.2e-15;
Matches 108; Conservative 70; Mismatches 170; Indels 82; Gaps 19;
QY 51 IMPDAGSTGRHIV--TFVQKMPGQLPIL-----ZGEVFD--SVKPGLSAFV 94
DB 525 LVTVISITGRATVYQASINYYKDDSLPIVWKSLETSRKSRCRAYDMETEPGPKLV 594
QY 95 DQPKQGAET-VOGLLEVAKDSIPRSHWKTTPVVLKATAGLRLPEHKAALLFEVKEIFR 153
DB 585 NN-RTGLKTAIKELIQAWEKQIPKNAHRTTSLFVATAGVRRLRPADSSWILGNWSILA 643
QY 154 KSPFLVPKGSVSIWTCODEGIFANVTNFTLGTOLHG-HRQETVGTLDGASTQITFLPQ 212
DB 644 KSPFTRCRREWKLIISGTEAYFCWALNYQTSMLGALPKATFGALDGLGSSLQVTFENE 703
QY 213 FEXTLEQTTPRGYLTSPFMNSTYKLYTHSYLGFGLKAA-----RLATL----- 255
DB 704 -ERTHNETN---LNLRIQSVNHLGAYSAGYGLNDAPDRSVVHLKLPVWKSDDLIE 758
QY 256 GALETGTGHTFRSACLPRWLAEWIFPGVKYQYQGNQGEVGFPE-----PCYAEVLVR 309
DB 759 GKLE-----MKHPCLSNGYNGYICQSCASSVQGGKKGSGVSIKLVGAPNNGECSA 810
QY 310 VVR-GKXHOPEEVR--GSPYAFSYVDRAVDTDMDYBKGGI--LKVEDPERKAREVC 363
DB 811 LAKNARCALPDGYPRHGQYAVSGPF-----VYRPFNLASASLDDVLEKREFC 862
QY 364 DN-----LENPTSGSPFL---CMDLSYITALLKDGFGFADSTV-----LQITKKVNIET 410
DB 863 DKAWQVARTSVSPQPTIEQYCFRAPYIVSLREGLYITDKQIIIGSGSITWLGVALLES 922
QY 411 GWALGATFHL 420
DB 923 GKALSITLGL 932

RESULT 15
C86276
7A19.33 protein - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 18-Nov-2002
C;Accession: C86276
R;Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.P.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.; Li, J.H.; Li, Y.; Lin, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: C86276
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-508 <STO>
A;Cross-references: GB:AE005172; NID:G5080800; PIDN:AAD39310.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1
C;Superfamily: nucleoside triphosphatase chromatin-associated

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2004, 13:42:31 ; Search time 10 Seconds
(without alignments)
2228.603 Million cell updates/sec

Title: US-10-091-085-3

Perfect score: 2250

Sequence: 1 MATSWGVTFMLVWVCVCSA.....ETGVALGATFHLQSLGISH 428

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2250	100.0	428	1 ENP5 HUMAN	Q75356 homo sapien
2	1994.5	88.6	427	1 ENP5 MOUSE	Q9wuz9 mus musculus
3	1933.5	85.9	429	1 ENP5 MESAU	Q9qyc8 mesocricetu
4	999	44.4	484	1 ENP6 HUMAN	Q75354 homo sapien
5	986	43.8	455	1 ENP6 RAT	Q9er31 rattus norv
6	489	21.7	455	1 NTPA PEA	P52914 pisum sativ
7	483.5	21.5	518	1 GDAL YEAST	P32621 saccharomyc
8	446.5	19.8	454	1 APY SOITU	P80595 solanum tub
9	387.5	17.2	485	1 YB4 CAEEL	Q18411 caenorhabdi
10	385	17.1	552	1 YB4 CAEEL	Q21815 caenorhabdi
11	383	17.0	493	1 ENP1 CHICK	Q93295 gallus gall
12	372	16.5	630	1 YND1 YEAST	P40009 saccharomyc
13	370	16.4	510	1 ENP1 MOUSE	P55772 mus musculu
14	365	16.2	513	1 ENP1 BOVIN	Q18956 bos taurus
15	360.5	16.0	613	1 ENP4 MOUSE	Q9dbt4 mus musculu
16	357	15.9	616	1 ENP4 HUMAN	Q9y2b7 homo sapien
17	352	15.6	510	1 ENP1 PIG	Q9myu4 sus scrofa
18	351	15.6	510	1 ENP1 HUMAN	P49561 homo sapien
19	348.5	15.5	511	1 ENP1 RAT	P97687 rattus norv
20	347.5	15.4	529	1 ENP3 HUMAN	Q75355 homo sapien
21	339.5	15.1	495	1 ENP2 HUMAN	Q9y213 homo sapien
22	309.5	13.8	495	1 ENP2 MOUSE	O55026 mus musculu
23	302.5	13.4	495	1 ENP2 RAT	Q35795 rattus norv
24	301.5	13.4	494	1 ENP2 CHICK	P79784 gallus gall
25	140	6.2	628	1 NTP1 TOXGO	Q27893 toxoplasma
26	138	6.1	628	1 NTP2 TOXGO	P52913 toxoplasma
27	131	5.8	592	1 NTP4 TOXGO	Q46455 mcoerella th
28	111.5	5.0	634	1 SELB MOOTH	P13333 spiroplasma
29	106.5	4.7	553	1 VGL SPV4	Q05470 bacillus su
30	104.5	4.6	4427	1 PKSL BACSU	O83273 treponema p
31	101.5	4.5	1151	1 Y245 TREPA	Q54796 streptococc
32	99.5	4.4	535	1 DEXB STRPN	P06874 bacillus st
33	99.5	4.4	548	1 THER_BACST	

34	97	4.3	1220	1	PTC1 PRARE	Q98864 brachydanio
35	96	4.3	711	1	DDP3 YEAST	Q08225 saccharomyc
36	95.5	4.2	2076	1	PAS1_YARLI	P34229 Y fatty aci
37	93	4.1	969	1	PAC4_HUMAN	P29122 homo sapien
38	93	4.1	1696	1	ITM2_HUMAN	Q9nm33 homo sapien
39	92	4.1	290	1	HEM6_CAUCR	Q9ast8 caulobacter
40	91	4.0	417	1	GLYA_BORBU	O51547 borrelia bu
41	91	4.0	477	1	URT1_DESSO	P98119 desmodus ro
42	90.5	4.0	301	1	PSD_CHLTR	O84705 chlamydia t
43	90	4.0	434	1	YFEW_ECO57	Q8xbj0 escherichia
44	90	4.0	434	1	YFEW_ECOLI	P77619 escherichia
45	89.5	4.0	601	1	SYR_GLOVI	Q7ndf6 gloeobacter

ALIGNMENTS

RESULT 1						
ENP5 HUMAN						
ID	ENP5 HUMAN	STANDARD;	PRT;	428 AA.		
AC	O75356;					
DT	16-OCT-2001 (Rel. 40, Created)					
DT	16-OCT-2001 (Rel. 40, Last sequence update)					
DT	15-MAR-2004 (Rel. 43, Last annotation update)					
DE	Ectonucleoside triphosphate diphosphohydrolase 5 precursor					
DE	(EC 3.6.1.6) (NTPDase5) (Nucleoside diphosphatase) (CD39 antigen-like					
DE	4) (ER-JDPase).					
GN	ENTPD5 OR CD39L4.					
OS	Homo sapiens (Human).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
OX	NCBI_TaxID=9606;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	TISSUE=Leukemia;					
RX	MEDLINE=98341119; PubMed=9676430;					
RA	Chadwick B.P., Frischauf A.-M.;					
RT	"The CD39-like gene family: Identification of three new human members					
RT	(CD39L2, CD39L3, and CD39L4), their murine homologues, and a member of					
RT	the gene family from Drosophila melanogaster."					
RL	Genomics 50:357-367(1998).					
CC	FUNCTION: Likely to promote reglycosylation reactions involved in					
CC	glycoproteins folding and quality control in the endoplasmic					
CC	reticulum. Hydrolyzes UDP, GDP AND IDP but not any other					
CC	nucleoside di-, mono- or triphosphates, nor thiamine pyrophosphate					
CC	(By similarity).					
CC	!- CATALYTIC ACTIVITY: A nucleoside diphosphate + H(2)O = a					
CC	nucleotide + phosphate.					
CC	!- COFACTOR: Requires calcium and magnesium.					
CC	!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen (By similarity).					
CC	!- TISSUE SPECIFICITY: Expressed in adult liver, kidney, prostate,					
CC	testis and colon. Much weaker expression in other tissues.					
CC	!- SIMILARITY: Belongs to the GDAL / CD39 NTPase family.					
CC	-----					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -					
CC	the European Bioinformatics Institute. There are no restrictions on its					
CC	use by non-profit institutions as long as its content is in no way					
CC	modified and this statement is not removed. Usage by and for commercial					
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/					
CC	or send an email to license@isb-sib.ch).					
CC	-----					
DR	EMBL; AF039918; AAC39885.1; -					
DR	Genew; HGNC:3367; ENTPD5.					
DR	MIM; 603162; -					
DR	InterPro; IPR000407; GDAL_CD39_NTPase.					
DR	Pfam; PF01150; GDAL_CD39; 1.					
DR	PROSITE; PS01236; GDAL_CD39_NTPASE; FALSE NEG.					
KW	Hydrolase; Transmembrane; Glycoprotein; Calcium; Magnesium;					
KW	Endoplasmic reticulum; Signal.					
FT	SIGNAL 1 20					
FT	CHAIN 21 428					
FT	ECTONUCLEOSIDE TRIPHOSPHATE					
FT	DIPHOSPHOHYDROLASE 5.					


```
FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 232 232 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 428 AA; 47517 MW; 830437A155DE4DDDD CRC64;

Query Match 100.0%; Score 2250; DB 1; Length 428;
Best Local Similarity 100.0%; Pred. No. 6.9e-172;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATSGTGVFFMLVWSCVCSAVSHRNQQTWFGIFLSSMCPINVSASTLYGIMFAGSGTGT 60
Db 1 MATSGTGVFFMLVWSCVCSAVSHRNQQTWFGIFLSSMCPINVSASTLYGIMFAGSGTGT 60
QY 61 RIHVTVFVQKPGQQLPILGEVFDVSVKPGCLGAFVDQPKQAGETVQGLLEVAKDISPRSHW 120
Db 61 RIHVTVFVQKPGQQLPILGEVFDVSVKPGCLGAFVDQPKQAGETVQGLLEVAKDISPRSHW 120
QY 121 KKTVPVLKATAGLRLLPHEKAKALLFEVKEIFRKSFPFLVPKGSVIMDGSDEGILAVTV 180
Db 121 KKTVPVLKATAGLRLLPHEKAKALLFEVKEIFRKSFPFLVPKGSVIMDGSDEGILAVTV 180
QY 181 NFLTQLGKHQETVGTLDLGASQTITFLQPKETLEQTTPRGVLTSPFNSTYKLYTH 240
Db 181 NFLTQLGKHQETVGTLDLGASQTITFLQPKETLEQTTPRGVLTSPFNSTYKLYTH 240
QY 241 SYLGFGLKAARLATIGALETETDGTHTFRSACLPRWLEAEWIFGVKYQYQGNQEGVG 300
Db 241 SYLGFGLKAARLATIGALETETDGTHTFRSACLPRWLEAEWIFGVKYQYQGNQEGVG 300
QY 301 EPCYAEVLVRVGRKLHQPVEVQSGFYAFSYYYDRAVDITMDIYKGGILKVEDFERKAR 360
Db 301 EPCYAEVLVRVGRKLHQPVEVQSGFYAFSYYYDRAVDITMDIYKGGILKVEDFERKAR 360
QY 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGPADSTVLIQTKKVNNIETGALGATFHL 420
Db 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGPADSTVLIQTKKVNNIETGALGATFHL 420
QY 421 LQSLGISH 428
Db 421 LQSLGISH 428

RESULT 2
ID ENP5 MOUSE STANDARD; PRT; 427 AA.
AC Q9WU29; 070214;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 5 precursor
DE (EC 3.6.1.6) (NTPDase5) (Nucleoside diphosphatase) (CD39 antigen-like
DE 4) (ER-UDPase).
GN ENTPD5 OR CD39L4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=98113025; PubMed=9457681;
RA Chadwick B.P., Williamson J., Sheer D., Frischauf A.-M.;
RT "cDNA cloning and chromosomal mapping of a mouse gene with homology to
RT NTPases."
RL Mamm. Genome 9:162-164(1998).
[2]
RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND CHARACTERIZATION.
RC TISSUE=Liver;
RX MEDLINE=9929181; PubMed=10369669;
RA Trombetta E.S., Helenius A.;
RT "Glycoprotein reglucosylation and nucleotide sugar utilization in the
RT secretory pathway: identification of a nucleoside diphosphatase in the
RT endoplasmic reticulum.";
```

```
RL EMBO J. 18:3282-3292(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Iwata M., Fushimi Y., Kondo H., Adachi J., Fukuda S.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Schirni L.M., Scubli R., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayaishizaki Y.;
RL Nature 409:685-690(2001).
[4]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Dratchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loughlan N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Morley D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Faney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smaluk D.E.,
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Likely to promote reglucosylation reactions involved in
CC glycoproteins folding and quality control in the endoplasmic
CC reticulum. Hydrolyzes UDP, GDP AND IDP but not any other
CC nucleoside di-, mono- or triphosphates, nor thiamine
CC pyrophosphate.
CC -!- CATALYTIC ACTIVITY: A nucleoside diphosphate + H(2)O = a
CC nucleotide + phosphate.
CC -!- COFACTOR: Requires calcium and magnesium.
CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC -!- TISSUE SPECIFICITY: Ubiquitous.
CC -!- PTM: Glycosylated with high mannose N-linked glycans.
CC -!- MISCELLANEOUS: Optimal pH is neutral.
CC -!- SIMILARITY: Belongs to the GDAI / CD39 NTPase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF006482; AAC05181.1; --
DR EMBL; AJ238636; CAB45533.1; --
```

```

DR EMBL; AK002618; BAB22234.1; -.
DR EMBL; BC015247; AAL15247.1; -.
DR MGD; MGI:1321385; Entp45.
DR InterPro; IPR000407; GDA1_CD39_NTPase.
DR Pfam; PF01150; GDA1_CD39; 1.
DR PROSITE; PS01238; GDA1_CD39_NTPASE; FALSE NEG.
KW Hydrolase; Transmembrane; Glycoprotein; Calcium; Magnesium;
KW Endoplasmic reticulum; Signal.
FT SIGNAL 1 18
FT CHAIN 19 427
FT CATALYTIC ACTIVITY: A nucleoside diphosphate + H(2)O = a
FT EC: 3.6.1.6 (NTPDase)
FT NUCLEOSIDE DIPHOSPHATE
FT DIHYDROPHOSPHATE
FT CARBOHYD 41 41
FT CARBOHYD 231 231
FT CONFLICT 390 390
FT CONFLICT 394 427
FT PLAKESORHRRDWGLGHLSPAPVSGHQLRPSSTSRAC
FT ISEVFSQEGVDSSTFSDSGKAWPSTR (IN REF. 1).
FT SQ SEQUENCE 427 AA; 47101 MW; 2F9DA2C342C5577 CRC64;

Query Match 88.6%; Score 1994.5; DB 1; Length 427;
Best Local Similarity 88.3%; Pred. No. 1.6e-151;
Matches 377; Conservative 24; Mismatches 25; Indels 1; Gaps 1;

QY 1 MATSWGTVPFVPMVAVSVCSAVSHRNQOTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
DB 1 MATSWGAV-FMLIIACVGVSTVFYREQOTWPEGVFLSSMCPINVSAGTLYGIMFDAGSTGT 59
QY 61 RIHVYTFVQKMPQLPILGEVDFSVKGLSAFVDQPKGAETVQGLLEVAKDSIPRSHW 120
DB 60 RIHVYTFVQKMPQLPILGEVDFSVKGLSAFVDQPKGAETVQGLLEVAKDSIPRSHW 119
QY 121 KETPVVLKATAGIRLLPEHAKALLFEVKEIFRKSPPFLVPGKSVSMDGSDGILAVTV 180
DB 120 ERTPVVLKATAGIRLLPEHAKALLFEVKEIFRKSPPFLVPGKSVSMDGSDGILAVTV 179
QY 181 NFLTGQLHGHROBTWGTLDLGGASTQITFLPQFKTLQTPRGYLTSPFENSTPKLYTH 240
DB 180 NFLTGQLHGHROBTWGTLDLGGASTQITFLPQFKTLQTPRGYLTSPFENSTPKLYTH 239
QY 241 SYLGFGKLAARLATIAGALETEGDTGHTFRSACLPRWLEAWIFGGVKYQYGGNQEVEGF 300
DB 240 SYLGFGKLAARLATIAGALETEGDTGHTFRSACLPRWLEAWIFGGVKYQYGGNQEVEGF 299
QY 301 EPCYAEVLVRVKGKLOPEVQSGSFYAFSYVDRAVDTMDIDYKGGILKVEDFERKAR 360
DB 300 EPCYAEVLVRVQKLOPEVQSGSFYAFSYVDRAVDTMDIDYKGGILKVEDFERKAR 359
QY 361 EVCNDLENFTSGSPFLCNDLSYITALLKQGFADSTVLQLTKKVNNIETGALGATEHL 420
DB 360 EVCNDLGSFSGSGSFFLCNDLTYITALLKQGFADSTVLQLTKKVNNIETGALGATEHL 419
QY 421 LQSLIGIS 427
DB 420 LQSLIGIT 426

RESULT 3
ID ENPS MESAU STANDARD; PRT; 469 AA.
AC Q9QYCB; Q9QYCB;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 5 precursor
DE (EC 3.6.1.6) (NTPDases) (Nucleoside diphosphatase) (CD39 antigen-like
DE 4) (ER-UDPase) (Proto-oncogene cph).
GN ENTPD5 OR CD39L4 OR CPH.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]

```

```

RP MEDLINE; 99142925; PubMed-9989819;
RA Velasco J.A., Avila M.A., Notario V.;
RT "The product of the cph oncogene is a truncated, nucleotide-binding
RT protein that enhances cellular survival to stress.";
RL Oncogene 18:689-701(1999).
CC -!- FUNCTION: Likely to promote reglycosylation reactions involved in
CC glycoproteins folding and quality control in the endoplasmic
CC reticulum. Hydrolyzes UDP, GDP AND IDP but not any other
CC nucleoside di-, mono- or triphosphates, nor thiamine pyrophosphate
CC (By similarity).
CC -!- CATALYTIC ACTIVITY: A nucleoside diphosphate + H(2)O = a
CC nucleotide + phosphate.
CC -!- COFACTOR: Requires calcium and magnesium (By similarity).
CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen (By similarity).
CC -!- TISSUE SPECIFICITY: Expressed in fetal cells and most adult
CC tissues.
CC -!- DISEASE: Cph oncogene has transforming capacity and tumorigenic
CC potential.
CC -!- SIMILARITY: Belongs to the GDA1 / CD39 NTPase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF084568; AAF22931.1; GDA1_CD39_NTPase.
CC InterPro; IPR000407; GDA1_CD39_NTPase.
CC Pfam; PF01150; GDA1_CD39; 1.
CC PROSITE; PS01238; GDA1_CD39_NTPASE; FALSE NEG.
KW Hydrolase; Transmembrane; Glycoprotein; Calcium; Magnesium;
KW Endoplasmic reticulum; Signal; Proto-oncogene.
FT SIGNAL 1 18
FT CHAIN 19 469
FT CATALYTIC ACTIVITY: A nucleoside diphosphate + H(2)O = a
FT EC: 3.6.1.6 (NTPDase)
FT NUCLEOSIDE DIPHOSPHATE
FT DIHYDROPHOSPHATE
FT CARBOHYD 42 42
FT CARBOHYD 232 232
FT SEQUENCE 469 AA; 52125 MW; 03D8A23E0C73474B CRC64;

Query Match 85.9%; Score 1933.5; DB 1; Length 469;
Best Local Similarity 86.7%; Pred. No. 1.4e-146;
Matches 364; Conservative 22; Mismatches 31; Indels 3; Gaps 1;

QY 1 MATSWGTVPFVPMVAVSVCSAVSHRNQOTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
DB 1 MATPWGAVFPLLMIAICAGSTVFYREQOTWPEGVFLSSMCPANVSASTLYGIMFDAGSTGT 60
QY 61 RIHVYTFVQKMPQLPILGEVDFSVKGLSAFVDQPKGAETVQGLLEVAKDSIPRSHW 120
DB 61 RIHVYTFVQKMPQLPILGEVDFSVKGLSAFVDQPKGAETVQGLLEVAKDSIPRSHW 120
QY 121 KETPVVLKATAGIRLLPEHAKALLFEVKEIFRKSPPFLVPGKSVSMDGSDGILAVTV 180
DB 121 KETPVVLKATAGIRLLPEHAKALLFEVKEIFRKSPPFLVPGKSVSMDGSDGILAVTV 180
QY 181 NFLTGQLHGHROBTWGTLDLGGASTQITFLPQFKTLQTPRGYLTSPFENSTPKLYTH 240
DB 181 NFLTGQLHGHROBTWGTLDLGGASTQITFLPQFKTLQTPRGYLTSPFENSTPKLYTH 240
QY 241 SYLGFGKLAARLATIAGALETEGDTGHTFRSACLPRWLEAWIFGGVKYQYGGNQEVEGF 300
DB 241 SYLGFGKLAARLATIAGALETEGDTGHTFRSACLPRWLEAWIFGGVKYQYGGNQEVEGF 300
QY 301 EPCYAEVLVRVKGKLOPEVQSGSFYAFSYVDRAVDTMDIDYKGGILKVEDFERKAR 360
DB 301 EPCYAEVLVRVQKLOPEVQSGSFYAFSYVDRAVDTMDIDYKGGILKVEDFERKAR 360
QY 361 EVCNDLENFTSGSPFLCNDLSYITALLKQGFADSTVLQLTKKVNNIETGALGATEHL 420
DB 361 EVCNDLGSFSGSGSFFLCNDLTYITALLKQGFADSTVLQLTKKVNNIETGALGATEHL 419

```


RP SEQUENCE FROM N.A.; SUBCELLULAR LOCATION, AND CHARACTERIZATION.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=20498744; PubMed=11042118;
RA Braun N., Fengler S., Ebeling C., Servos J., Zimmermann H.;
RT "Sequencing, functional expression and characterization of rat
RT NTPase6, a nucleoside diphosphatase and novel member of the ecto-
RT nucleoside triphosphate diphosphohydrolase family.";
RL Biochem. J. 351:639-647(2000).
CC -!- FUNCTION: Might support glycosylation reactions in the Golgi
CC apparatus and, when released from cells, might catalyze the
CC hydrolysis of extracellular nucleotides. Hydrolyzes preferentially
CC nucleoside 5'-diphosphates, nucleoside 5'-triphosphates are
CC hydrolyzed only to a minor extent, there is no hydrolysis of
CC nucleoside 5'-monophosphates. The order of activity with different
CC substrates is GDP > IDP > UDP = CDP >> ADP.
CC -!- CATALYTIC ACTIVITY: A nucleoside diphosphate + H(2)O = a
CC nucleotide + phosphate.
CC -!- COFACTOR: Requires calcium and magnesium.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. GOLGI. BUT ALSO
CC OCCURS IN A SOLUBLE EXTRACELLULAR FORM.
CC -!- SIMILARITY: Belongs to the GDA1 / CD39 NTPase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
CC or send an email to license@isb-sib.ch.
CC
CC EMBL: AJ277748; CAC16598.1; -
CC InterPro: IPR000407; GDA1_CD39_NTPase.
CC Pfam: PF01150; GDA1_CD39; 1.
CC PROSITE: PS01238; GDA1_CD39_NTPASE; FALSE NEG.
CC Hydrolase; Transmembrane; Glycoprotein; Calcium; Magnesium;
CC Signal-anchor; Golgi stack.
CC DOMAIN 1 12 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 13 32 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
CC (POTENTIAL).
CC FT DOMAIN 33 455 LUMENAL (POTENTIAL).
CC FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 256 256 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SQ SEQUENCE 455 AA; 49899 MW; 19A32E8BAEF0F77B CRC64;
Query Match 43.8%; Score 986; DB 1; Length 455;
Best Local Similarity 54.0%; Pred. No. 4.7e-71; Indels 6; Gaps 5;
Matches 204; Conservative 51; Mismatches 117;
Qy 49 YGIMFDAGSTGTRIHYVTFVQKMPGLPILLEGVFDSPVQKGLSAFVDPKQGAETVQGLL 108
Db 74 YGIMFDAGSTGTRIHYVQFPA-RPPGETPTLTHTETFKALKPGLSAYADDDVEKSAQGIQELL 132
Qy 109 EVAKDSIPRSHWKKTPVVLKATAGRLLLPEHKAKALLFEVKEIF-RKSPFLVPKGSVSI 168
Db 133 NVAKQHIPPYDFWKATPLVLKATAGRLLLPGEKQKLLQKRVFKASPLFLVGDDCVSIWN 192
Qy 169 GSDEGILAWTVNPLTQGLHGHQETVTGLDLGASQTITFLPQPEKTLTQTPRGYITSP 228
Db 193 GTDEGVSAWITVNLFTGLSLTQPGSSVGMGLDGGSTQITFLPRVEGTQLQASPGHULTAL 252
Qy 229 EMFNSTKLTHTSYLGLFGLKAARLATLGALETB-CTDGTHTFRACILPRMLEAWIEGGYK 287
Db 253 QMFNFTPLYSYSLYGLGLMSARLATLGGYEGEPABDDKELVSPCLSPFRGKWEHART 312
Qy 288 YQYGEQGEVGVG-PEPCYAEVLVRVKGHLKHOPEEVQVGSFYAFSYYYDRAVDTDMDIDYEK 346
Db 313 YRISGQK--AVGLYELCASRVSEVLNKHVHRTSEAQVDFYAFSYYYDLAASFGLDAEK 370
Qy 347 GGILKVEDFERKAREVCNLENTSSGPFLLCMDSLYITALLKDGFGFADSTVLQTKYKN 406
Db 371 GGSILVGDFFBIAAKYVCKTLETOPSPSPACMDLTVISLLHE-FGPPGDKVLKARKID 429
Qy 407 NIETCWALGATFHLQLSL 424

Db 430 NVETSWALGAIFHYIDSL 447
RESULT 6
NTPA_PEA
ID NTPA_PEA STANDARD; PRT; 455 AA.
AC P52914;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Nucleoside-triphosphatase (EC 3.6.1.15) (Nucleoside triphosphate
DE diphosphohydrolase) (NTPase) (Apyrase).
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
OX NCBI_TaxID=3888;
RX [1]
RY SEQUENCE FROM N.A.
RP STRAIN=cnv. Alaska; TISSUE=Plumule;
RX MEDLINE=96197404; PubMed=8616230;
RA Hsieh H., Tong C.G., Thomas C., Roux S.J.;
RT "Light-modulated abundance of an mRNA encoding a calmodulin-regulated,
RT chromatin-associated NTPase in pea.";
RL Plant Mol. Biol. 30:135-147(1996).
RN [2]
RY SEQUENCE FROM N.A.
RP STRAIN=cnv. Alaska; TISSUE=Stem;
RA Shibata K., Abe S., Davies E.;
RT "Structure of the coding region and mRNA variants of the apyrase from
RT Pisum sativum.";
RL Acta Physiol. Plant. 20:3-13(2001).
CC -!- FUNCTION: Might be involved in RNA transport out of nuclei.
CC -!- CATALYTIC ACTIVITY: NTP + H(2)O = NDP + phosphate.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the GDA1 / CD39 NTPase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
CC or send an email to license@isb-sib.ch.
CC
CC EMBL: Z32743; CAA83655.1; -
CC EMBL: AB027614; BAB18900.1; -
CC PIR: S65147; S48859.
CC InterPro: IPR000407; GDA1_CD39_NTPase.
CC Pfam: PF01150; GDA1_CD39; 1.
CC PROSITE: PS01238; GDA1_CD39_NTPASE; 1.
CC Hydrolase; Nuclear protein.
CC SQ SEQUENCE 455 AA; 50072 MW; 50FDF0023ABC4299 CRC64;
Query Match 21.7%; Score 489; DB 1; Length 455;
Best Local Similarity 33.7%; Pred. No. 2e-31;
Matches 140; Conservative 62; Mismatches 160; Indels 54; Gaps 14;
Qy 49 YGIMFDAGSTGTRIHYVTFVQKMPGLPILLEG-EVFDSPVKPGLSAFVDPKQGAETVQGL 107
Db 44 YAVVFDAGSTGTRIHYVTFVQKMPGLPILLEG-EVFDSPVKPGLSAFVDPKQGAETVQGL 102
Qy 108 LEVAKDSIPRSHWKKTPVVLKATAGRLLLPEHKAKALLFEVKEIF-RKSPFLVPKGSVSI 166
Db 103 LEQAEVVPDDLPQKTPVRLGATAGRLLLNGDASSKILQSVRDMLSNRSTFNVQPDVSI 162
Qy 167 MDGSGEGILAWTVNPLTQGLHGHQETVTGLDLGASQTITFLPQPEKTLTQTPR- 222
Db 163 IDGTGEGSVLWTVNPLTQGLHGHQETVTGLDLGASQTITFLPQPEKTLTQTPR- 221
Qy 223 --GYLTSPFEMENSTYKLYTHSYLGLFGLKAARLATLGALETGDTGHTFES--ACPRWLE 278

Db 222 DDPIKVKVLLGIPYDLYVHSLYHFGREASRAEILKL-----TPRSPNPLLAFGN 272
QY 279 AEWIFGVYKVOYGGNVEGVFPCEVAEVLVVRGKLHOPVEVQGRSP----- 326
Db 273 GIYVSGEGRFAPATYSG-ANFNKCKNTIRKAL--KLNPYCPYQNTCFGIWNGGGNGQ 329
QY 327 ---YAFSYYYDRAVDTMDIDYKGG-ILKVEDEPKAREVCD-NLENPTSGSPFL----- 376
Db 330 KNLFASSSPFLPDGWDASTPFIIRPVDIETRAKEACALNFDASTYFPLDKKV 389
QY 377 ----CWLSTYITALLKOGFGPADSTVLQTKKVN-----IEFGWALGATFHLLQSL 424
Db 390 ASYVCMDLIYQYVLLVDFGLDPLQKTSKTEIYQDAIVEAAMPIGNAVEAISAL 445

RESULT 7
GDAL_YEAST
ID GDAL_YEAST STANDARD; PRT; 518 AA.
AC F32621;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Guanosine-diphosphatase (EC 3.6.1.42) (GDPase).
GN GDAL OR YEL042N OR SYCP-ORF06.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=G2-9.
RX MEDLINE=93308137; PubMed=8391537;
RA Abeljon C., Yanagisawa K., Mandon E.C., Haessler A., Moremen K.,
RA Hirschberg C.B., Robbins P.W.;
RT "Guanosine diphosphatase is required for protein and sphingolipid
RT glycosylation in the Golgi lumen of Saccharomyces cerevisiae.";
RL J. Cell Biol. 122:307-323(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=97313264; PubMed=9169868;
RA Dietrich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,
RA Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,
RA Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,
RA Hunnicke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,
RA Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P.,
RA Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,
RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;
RT "The nucleotide sequence of Saccharomyces cerevisiae
RL Nature 387:78-81(1997).
CC -!- FUNCTION: After transfer of sugars to endogenous macromolecular
CC acceptors, the enzyme converts nucleoside diphosphates to
CC nucleoside monophosphates which in turn exit the Golgi lumen in
CC a coupled antiporter reaction, allowing entry of additional
CC nucleotide sugar from the cytosol.
CC -!- CATALYTIC ACTIVITY: GDP + H(2)O = GMP + phosphate.
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Golgi.
CC -!- SIMILARITY: Belongs to the GDAL / CD39 NTPase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L19560; AAA34656.1; -.
DR EMBL; U18779; AAB65000.1; -.
DR PIR; A40732; A40732.
DR GeneOnline; 139046; -.
DR SGD; S0000768; GDAL.

GO; GO:0004382; P:guanosine diphosphatase activity; IDA.
GO; GO:0045134; P:uridine diphosphatase activity; IDA.
DR InterPro; IPR000407; GDAL_CD39_NTPase.
DR Pfam; PF01150; GDAL_CD39; 1.
DR PROSITE; PS01238; GDAL_CD39_NTPASE; 1.
KW Hydrolase; Golgi stack; Glycoprotein; Transmembrane; Signal-anchor.
FT DOMAIN 1 9 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 10 24 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT DOMAIN 25 518 LUMENAL (POTENTIAL).
FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 280 280 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 335 335 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 518 AA; 56821 MW; 9A61182D5ED22ADF CRC64;
Query Match 21.5%; Score 483.5; DB 1; Length 518;
Best Local Similarity 32.8%; Pred. No. 6.7e-31;
Matches 142; Conservative 56; Mismatches 160; Indels 75; Gaps 15;
QY 49 YGIMFDAGSTGTRIHYVTF-VQKMPGQLPILGEVDFDSVKPLGSAFVDQPKQGAETVQGL 107
DB 93 YVIMIDAGSTGSRVHIYKFDVCTSP---FTLLDEKFDMLBPCLSFSDTDSVGAANSLDPL 149
QY 108 LEVAKDSIPSRSHWKKTPVVLKATAGLRLLPEHKAKALLFEYKEIFRKS-PFLVPRG-SVS 165
DB 150 LKVAANYVPIKARCTPVAVKATAGLRLLGDAKSKILLSAVRDHLEKDYPPFVVEGDGVS 209
QY 166 IMDGSEGLAWTVNPLTQGL--HGHRQETVGTDLGGASTQITFLPQFKTLEQTTPRG 223
DB 210 IMGDEEGVPWITTNVLLGNIGANGPKLPTAAVFDLGGSGSTQVFEPTPINEKMWDE 269
QY 224 YLTSEFENSTYKLYTHSYLQFLGKAAAR-----LATLGALETGDTGHTFRSAC 272
DB 270 HKPDLKFGDENYTLQPSHLGYLKEGRKNVNSVLVNALKDGKILKGDNTKTKHQLSPC 329
QY 273 LPRMLEA--EWI-----EGVKYQYGG-----NQSGVGFEPQVAVELR 309
DB 330 LPPKVNATNEKVLKESKETYITDIPDEPSGAQCRFLTDILNKDAQCSPPCSF---- 385
QY 310 VVRGKLHOPEV-----ORGSYAFSYTYDRAVDOTDM-IDYKGGILKVEDFERKAVCV- 363
DB 386 ---NGVQPSLVRTFKESNDIYFSDYDRTPRLGMLPSF-----TLNELNDLARIYCK 436
QY 364 -----DNLENFTSGSPFLCWLSTYITALLKOGFGPADSTVLQTKKVNLET 410
DB 437 GEETNSVFSGIAGSLDELSDSHF-CLDISPQVSLHTGYDIPLORELTRGKLIANKEI 495
QY 411 GWALGATFHLLQOS 423
DB 496 GWCLGASLPLILKA 508

RESULT 8
APY_SOLITU
ID APY_SOLITU STANDARD; PRT; 454 AA.
AC P80595; Q43164;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Apyrase precursor (EC 3.6.1.5) (ATP-diphosphatase) (Adenosine
DE diphosphatase) (ADPase) (ATP-diphosphohydrolase).
GN RROPI.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 59-160; 236-253 AND 332-345.
RC TISSUE=Tuber;
RX MEDLINE=96158985; PubMed=8579614;
RA Handa M., Guidotti G.;
RT "Purification and cloning of a soluble ATP-diphosphohydrolase

(aprase) from potato tubers (Solanum tuberosum).";
 Biochem. Biophys. Res. Commun. 218:916-923(1996).
 [2]
 RN SEQUENCE OF 42-54; 68-95 AND 236-253.
 RC STRAIN-cv. Desires;
 RX MEDLINE=96355615; PubMed=8703025;
 RA Vasconcelos E.G., Ferreira S.T., de Carvalho T.M.U., de Souza W.,
 RA Kettum A.M., Mancilla M.A., Verjovski-Almeida S.,
 RT "Partial purification and immunohistochemical localization of ATP
 RT diphosphohydrolase from Schistosoma mansoni. Immunological cross-
 RT reactivities with potato apyrase and Toxoplasma gondii nucleoside
 RT triphosphate hydrolase";
 RL J. Biol. Chem. 271:22139-22145 (1996).
 CC -!- FUNCTION: Catalyzes the hydrolysis of phosphoanhydride bonds of
 CC nucleoside tri- and di-phosphates.
 CC -!- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 phosphate.
 CC -!- COFACTOR: Calcium.
 CC -!- SUBCELLULAR LOCATION: Membrane-associated (Probable).
 CC -!- PTM: The N-terminus is blocked.
 CC -!- SIMILARITY: Belongs to the GDAI / CD39 NTPase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U58597; AAB02720.1; -;
 DR PIR; J4616; J4616.
 DR InterPro; IPR000407; GDAI_CD39_NTPase.
 DR Pfam; PF01150; GDAI_CD39.1.
 DR PROSITE; PS01238; GDAI_CD39_NTPASE; 1.
 KW Hydrolase; Transmembrane; Calcium; Signal.
 FT SIGNAL 1 30 POTENTIAL.
 FT CHAIN 31 454 APYRASE.
 FT TRANSMEM 426 446 POTENTIAL.
 FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 454 AA; 50041 MW; 9D9EPE431D2A2F52F CRC64;
 Query Match 19.8%; Score 446.5; DB 1; Length 454;
 Best Local Similarity 28.8%; Pred. No. 5e-28;
 Matches 131; Conservative 78; Mismatches 177; Indels 69; Gaps 16;
 QY 25 NQOTWPGIFLS-----SMCPINYS-----STLYGMFDAGSGTGRHVV 65
 DB 3 NQNSHFIFILALFLVPLSLKKNVNAQIPLRHLLSHSEHYAVIFDAGSGSRVVF 62
 QY 66 TFVQXMPGQPILEG-EVFDVSVKPLSAFYDQPKQGAETVQGLLEVAKDSIPSHWKTTP 124
 DB 63 RFDEKL-GLLPIGNNIEYFNATEPGLSSVAEDPKAAANSLEPLDGAEGVWPQELQSETP 121
 QY 125 VVLKATAGLRLLPEHKAKALLPEVKEIFR-KSPFLVPKGSVSTMDGSDGILAWVTNVL 183
 DB 122 LELGATAGLRMLKGDAAEKLQAVNLVKNQSTFHSKQDQWITLDGTQEGSYNMAINYL 181
 QY 184 TGLHGRHQRQTVGTLDDGASTQTITFL---PQEKILE-OTPRGYLTSEFMFNSTYKLYT 239
 DB 182 LGNLGKDYKSTTATIDLGGSQVMAYAI SNEQAKAPQNEDEGEPYVQKHLMSKQVNLV 241
 QY 240 HSYLGPGLKARLATLGALETGTDGHTFRSACLPLWLEAWETFGYKIQYGGNQGEVGC 299
 DB 242 HSYLNYQLAGRAEIPKASRNES-----NPCALEGCDGYSGYGVYKVPKPKGS-- 292
 QY 300 FEPYAEVLAVVRG--KLHQPEEVORGSF-----YAFSVYVDRAVDTMDI 342
 DB 293 ---SWKCRRLUTHALKINAKNLEECTFNGVNGGGDQKNIHASSFFYDIGAQVIV 349
 QY 343 DYE-KGGILKVEDFERKAREVCD-NLENFTS-----GSPFLCMLDLSVITALLKDGFG 392
 DB 350 DTRFPSALAKPIQVLAANKVACQTNVADIKSIPFKQDRNIPYLCMDLIYEYTLVLDGFG 409

QY 393 F---ADSTVLQTKKVN-IEFGWALGATFHLQ 423
 DB 410 LNPHEKITVHDVQYKYLVGAAWPLGCAIDLVS 444
 RESULT 9
 Y4E_CABEL STANDARD; PRT: 485 AA.
 ID Y4E_CABEL
 AC Q18411;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 54.3 kDa protein C33H5.14 in chromosome IV.
 GN C33H5.14
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RA Bradshaw H., Stellyes L.;
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: Belongs to the GDAI / CD39 NTPase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U41007; AAA82272.1; -;
 DR PIR; T34147; T34147.
 DR WormPep; C33H5.14; CE04157.
 DR InterPro; IPR000407; GDAI_CD39_NTPase.
 DR Pfam; PF01150; GDAI_CD39.1.
 DR PROSITE; PS01238; GDAI_CD39_NTPASE; 1.
 KW Hypothetical protein; Transmembrane; Hydrolase.
 FT TRANSMEM 439 459 POTENTIAL.
 SQ SEQUENCE 485 AA; 54309 MW; DE64DIADC20F581E CRC64;
 Query Match 17.2%; Score 387.5; DB 1; Length 485;
 Best Local Similarity 27.7%; Pred. No. 2.8e-23;
 Matches 120; Conservative 84; Mismatches 152; Indels 77; Gaps 18;
 QY 49 YGIMFDAGSGTGRHVVTFVOKMPGQPILEGVFD-----VKPGLSAFVDQPKQGA 102
 DB 25 YGVICDAGSGTGRHVVTFVOKMPGQPILEGVFD-----VKPGLSAFVDQPKQGA 83
 QY 103 TVQGLLEVAKDSIPSHWKTTPVVLKATAGLRLLPEHKAKALLPEVKEIFRKSPPFL-VPK 161
 DB 84 YLTPLLRFAEEHIPYEQLGSETDLLIFATAGMELLPEAQKDAIKMLQNLKSVTLRVSD 143
 QY 162 GSVSMDGSDGILAWVTNFLTQGLHGRHQRQTVGTLDDGASTQTITELPQPEK----- 215
 DB 144 SNIRIDGAWEGHYKSIYAVNYTLGFRDENDSKVGMIDMGASVQIAFIAEKESYNGG 203
 QY 216 TLEQTPRGYLTSEFMFNSTYKLYTHSYLFGGLKAAARLATLGALETGTDGHTFRSACLPR 275
 DB 204 NYVEINLGSIEINEDYK--YKIYSTFLGYGANEGKLKYENSLVKSGNS-----NDSCSPR 257
 QY 276 WLEAEWIFGVYKQYGGNQGEVGCPEVAVLVRVVRGKLHQPE----- 319
 DB 258 GLNR--LIG-----EFTVNGTGE--MDVCLAQVSSLI-GDKAQPSPNFTCFLRNVIAVS 308
 QY 320 EYORGSTFAFS--YYIDRAVDTMDIYKGGILKVEDFERKAREVC-----DNLNFTSGSP 374
 DB 309 NLSTVQLGCFSEYWTTS-----NFGSGEYHYQKFTDEVKCYCKQKWNIDQDGFRNE 362
 QY 375 F-----LMDLSYITALLKDGFGFADST--VLQTKKVNIEFGWALG----- 416

Db 363 FPNADIERLGTNCFKAAWTSVLHDSFN-VDKTKHLFQSVLKIAGEEMQALGMILYHSK 421
Qy 417 --TFPHLLQSLIGIS 427
Db 422 DLKFNLEQLEVA 434
RESULT 10
YBU4_CABEL STANDARD; PRT; 552 AA.
AC Q21815;
DT 01-NOV-1997 (Rel. 35, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hydrothermal protein R0784.4 in chromosome X.
GN R0784.4
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
ON NCBI_TaxID=6239;
RX [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Bristol N2;
RC Miller N.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RA Waterston R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the GDA1 / CD39 NTPase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U39652; AAA80403.2;
DR WormPep; R0784.4; CE28748.
DR InterPro; IPR000407; GDA1_CD39_NTPase.
DR Pfam; PF01150; GDA1_CD39_1.
DR PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
KW Hypothetical protein; Transmembrane; Hydrolase.
FT TRANSMEM 7 27
FT TRANSNEM 490 510
FT POTENTIAL.
SQ SEQUENCE 552 AA; 62510 MW; B5DC32C858AE4D94 CRC64;
Query March 17.18; Score 385; DB 1; Length 552;
Best Local Similarity 28.68; Pred. No. 5.2e-23;
Matches 126; Conservative 69; Mismatches 160; Indels 86; Gaps 20;
Qy 49 YGINEDAGSTGTRIHYVTFVQKMPGQPLILEGEVFDs-----VKPLGSFVDPQKQAE 102
Db 44 YGVICDAGSTGTRLFVYVNWISTSDSELIQIEPVYDKNPKVWKISPLGSLTFGTPQAQAE 103
Qy 103 TVQGLLVAKDSIPRSHKKTTPVVLKATAGRLPEHKAKALLPEVKEIFRKSPEL---- 158
Db 104 YLRPLMEALERHPIEEKRPYPFPFATAGNRLIPDQKEAVL---KLNRLKLPKITSMQ 160
Qy 159 VPKGSVIMDGSDEGILAVTNFPLTGLH-----GH-RQETVGTLDLGAOST 205
Db 161 VLKEHRIIEGKMEGIYSWIAVNVALGFENKNTATLDFPGTSPAHARQKTVGMIDMGASA 220
Qy 206 QITP-LQPEK-----TLQETPRGVLTSFEMNSYKLYTHSYLGF-----GLKAAALATLG 256
Db 221 QIAELPDTDSFSSINVENIMLGREDSDLFK--YKLFTVTFLLGYGVNIGIRKYHEMLLS 278
Qy 257 ALETGTDGHTFRACLPRLWEAEWIFGVKYQXGKN--QEG-----E 297
Db 279 XLKQDQ--NGTVIQDDCNPLNHLK-----VTLENGENFVRRTGWNWTCNEVKKLNP 331

Qy 298 VGFPPCAEVLVRVVRGKHOPE-EVQRGSFYAFSVYYDRAVDTDMIDYKGGILKVEDFE 356
Db 332 SSSVEVCAEAAKCYFGAVPAPSPILSNIEVMYGPSEYVWSTHDV---LGLGGQYDAENIA 387
Qy 357 RKAREVCD-----NLENFTSGSP-----FLCMDSLYTALKKQGFSPADST--VLQ 400
Db 388 KKTQCYCKRWSTTQAEBSKQLYPRADDEERLRTQFKSAWTSVLHDSFN-VDKTHNKFQ 446
Qy 401 LTKKVNNTETGMALGA-TEHL 420
Db 447 SVSTIAGQEVQWALGMNIYEM 467
RESULT 11
ENP1_CHICK STANDARD; PRT; 493 AA.
AC O93295;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 1 (EC 3.6.1.5)
DE (NTPDase1) (Ecto-ATP diphosphohydrolase) (ATPDase) (Lymphoid cell
DE activation antigen) (Ecto-apyrase) (CD39 antigen).
GN ENTPD1 OR CD39.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-21 AND 150-156.
RC TISSUE=Oviduct;
RX MEDLINE=98298108; PubMed=9632655;
RA Nagy A.K., Knowles A.P., Nagami G.T.;
RT "Molecular cloning of the chicken oviduct ecto-ATP-
RT diphosphohydrolase."
RL J. Biol. Chem. 273:16043-16049 (1998).
RN [2]
RP SEQUENCE OF 1-17.
RC TISSUE=Stomach;
RX MEDLINE=97442428; PubMed=9295305;
RA Lewis-Carl S., Kirley T.L.;
RT "Immunolocalization of the ecto-ATPase and ecto-apyrase in chicken
RT gizzard and stomach. Purification and N-terminal sequence of the
RT stomach ecto-apyrase."
RL J. Biol. Chem. 272:23645-23652 (1997).
CC -1- FUNCTION: In the nervous system, could hydrolyze ATP and other
CC nucleotides to regulate purinergic neurotransmission. Could also
CC be implicated in the prevention of platelet aggregation.
CC Hydrolyzes ATP and ADP equally well (By similarity).
CC -1- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 phosphate.
CC -1- COFACTOR: Requires calcium and magnesium (By similarity).
CC -1- SUBUNIT: Homodimer; disulfide-linked (Probable).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- PTM: N-glycosylated.
CC -1- SIMILARITY: Belongs to the GDA1 / CD39 NTPase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF041355; AAC26491.1;
DR InterPro; IPR000407; GDA1_CD39_NTPase.
DR Pfam; PF01150; GDA1_CD39_1.
DR PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
KW Hydrolase; Transmembrane; Antigen; Glycoprotein; Calcium; Magnesium.
FT DOMAIN 1 7
FT TRANSMEM 8 28
FT POTENTIAL.

FT DOMAIN 29 463 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 464 486 POTENTIAL.
FT DOMAIN 487 493 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 65 65 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 73 73 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 133 133 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 223 223 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 267 267 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 324 324 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 330 330 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 361 361 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 372 372 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 382 382 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 16 16 I -> W (IN REF. 2).
FT CONFLICT 21 21 I -> G (IN REF. 2).
SQ SEQUENCE 493 AA; 54034 MW; F14FF4C3AA2F3603 CRC64;

Query Match 17.0%; Score 383; DB 1; Length 493;
Best Local Similarity 27.3%; Pred. No. 6.4e-230;
Matches 131; Conservative 66; Mismatches 170; Indels 112; Gaps 18;

QY 11 MLVSVCSGAVSHRNQQTWFGIFLSSMCPINVSASTLYGIMFDAGSTGTRIHVYTF-VQ 69
DB 11 LITATCVFSIIALLSANDVKDFL-----PPCTKGLNFDAGSTHATLYVYQMPAD 62

QY 70 KMPGQPILEGVDFDSVKGLSAFVDPQKQAGTVOGLLEVAKDSIPRSHWKTTPVVLKA 129
DB 63 KENGTGIVQVESCTVNGSGISSYADDPAGASLXPCLDKAMAVIPVBOQQTPTYLGA 122

QY 130 TAGRLRLPEH---KAKALLPEVKEIFRKPSPFLPKGSVIMDSGDEGILLAVTVNFTQ 186
DB 123 TAGRLRLRQNSTKARQVFAEVSXAIRFP--VDPRGAQLTGNESGFGNTVNYLLFT 180

QY 187 L-----HGHRQETVGLDGGASTQITPLPQFEKTLQTPRGYLTSPFEMFNSTYK 236
DB 181 LKFSFAGKWEHPQNTVELGALDGGASTQITPQGV--TIEDKNTSVL--FRLYGTNTS 236

QY 237 LYTHSVLGLCLXARLATLGALETECDGHTFESACLPRLWBAEWIFGKVKYQYGNQB- 295
DB 237 LYTHSVLGLYQIQASKRLMAALHQDGSYVQNIHPCYK-----GYRI 280

QY 296 ---GEVGFPEFY-----AEVLRYVR-----GKLLHQ 317
DB 281 ITIATVYDPCVTPSPMLSPAQLITVTGNPAACPTAILKFLNLTGCGANRTCGFDGVYQ 340

QY 318 PEEVQSGSYAFS-YYIDRA-----VDTMDYD-EKGGILKYEDPERKAREVC 363
DB 341 PP--VRGQFFAFAGFYTFESFLNLTCQQLSHVNATWDFCNKWSSELVETPFQNK---- 394

QY 364 DNLENSTGSPFLCMDLSYITALLKDGFGFADST--VLQLTQKVNNIETGWLGAFTHL 420
DB 395 EHLHYT-----CVVGLITLLVDGYKFDHTNSNIHFSQKAGNADIGTLGFMNL 446

RESULT 12
YND1_YEAST STANDARD; PRT; 630 AA.
AC P40009;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Golgi apyrase (EC 3.6.1.5) (ATP-diphosphatase) (Adenosine diphosphatase) (ADPase) (ATP-diphosphohydrolase) (Golgi nucleoside diphosphatase).
GN YND1 OR YER005W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.

RC STRAIN=S288c;
RX MEDLINE=99340091; PubMed=10409709;
RA Gao X.D., Kalgoredov V., Jigami Y.;
RT "YND1, a homologue of GDA1, encodes membrane-bound apyrase required for Golgi N- and O-glycosylation in Saccharomyces cerevisiae.";
RL J. Biol. Chem. 274:21450-21456(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=97313264; PubMed=9169868;
RA Dietrich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E., Araujo J., Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H., Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome V."; Nature 387:78-81(1997).
RL Nature 387:78-81(1997).
CC -!- FUNCTION: Catalyzes the hydrolysis of phosphoanhydride bonds of nucleoside tri- and di-phosphates. Has equal high activity toward ADP/ATP, GTP/GDP, and UDP/UTP and approximately 50% less toward CDP/CTP and thiamine pyrophosphate. Has no activity toward GMP. Required for Golgi glycosylation and cell wall integrity.
CC -!- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 phosphate.
CC -!- PATHWAY: Glycosylation.
CC -!- SUBCELLULAR LOCATION: Golgi.; membrane-bound.
CC -!- SIMILARITY: Belongs to the GDA1 / CD39 NTPase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; AF203695; AAF17573.1; --
CC EMBL; U18778; AAB64538.1; --
CC PIR; S50463; S50463.
CC GerMOnline; 139085; --
CC SGD; S0000807; YND1.
CC GO; GO:0017110; F:nucleoside diphosphatase activity; IDA.
CC GO; GO:0006486; P:protein amino acid glycosylation; IMP.
CC InterPro; IPR000407; GDA1_CD39_NTPase.
CC Pfam; PF01150; GDA1_CD39_1.
CC PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
KW Hydrolase; Transmembrane; Golgi stack.
FT DOMAIN 1 500 LUMENAL (POTENTIAL).
FT TRANSMEM 501 517 POTENTIAL.
FT DOMAIN 518 630 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 630 AA; 71851 MW; 02F8D24A78212544 CRC64;

Query Match 16.5%; Score 372; DB 1; Length 630;
Best Local Similarity 26.2%; Pred. No. 6.8e-22;
Matches 117; Conservative 73; Mismatches 165; Indels 92; Gaps 17;

QY 49 YGIMFDAGSTGTRIHVYTF-----VQMPQQLPILGE-----VFDSVVKPLGA 92
DB 10 FGIVIDAGSGSRHVFKNQDTSLLHATNQDSQILQSVPHIQEKDWTFLKPLGLSS 68

QY 93 PVDQPKQCAET-VQGLLEVAKDSIPRSHWKTTPVVLKATAGLRLPEHKAKALFEV-KE 150
DB 69 FEKKPQDQAYSHKIPLLDFARNIPESHWSGCPFIQTATAGRLLPDIQISILDLCCQ 128

QY 151 IFRKSPFLVPP--KGSVSIMDSGDEGILLAVTVNFTQLGHLGRQ-----TVGTLDLGA 203
DB 129 LKHPABELVEDCSAQIQVIDGETEGLYGMLGILNYLGHFNFDYNPEVSHFTFGFMDGGA 188

QY 204 STQITFLPQPEKTL-----QTPRGYLTSPFEMFNSTYKLYTHSYLGFGLKAARL 252
DB 189 STQIAFAPHSQGEIARHDDIATIFLRVNGDLQKQWDFVST-----WLGFGANQARR 241


```

QY 253 ATLGAL-----ETGCTDGTHTPSACLPRWLAEWIFGVKQYQVGNQGEVGP 301
Db 242 RYLAQLINTLPENTNDYENDDFSRNLNDFCMRGSSTDFEFDXTIFPHIAGSNYQCTK 301
QY 302 PCTAEVL-----VVRGKLHQP-EVORGFSFYAFSYTYDRAVTDMDYKGGILKV 352
Db 302 SIYPLLLKNPCDDPCLFNGVHAPRIDFANDXFGTSETSYWYANDV-----FKLGGEYNF 357
QY 353 EDPERKAREVCDN-----LENFTSG-----SPL-----CMDLSVITALLKDGFGFA---- 394
Db 358 DKSKSLREPCNSNWTOLANSKGVNSIPENFLKDACFKGNWLVNLHIEGFMORIDV 417
QY 395 -----DSTVLQLTKVNNIETGWALG 415
Db 418 DAENVNDRPLFSQVKEVKEBELSWTLG 444

RESULT 13
ENPI_MOUSE STANDARD; PRT; 510 AA.
AC P55772;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 1 (EC 3.6.1.5)
DE (NTPDase1) Ecto-ATP diphosphohydrolase (ATPDase) (lymphoid cell
DE activation antigen) (Ecto-apyrase) (CD39 antigen).
GN ENPPI OR CD39.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95015846; PubMed=7930580;
RA Maliszewski C.R., Deleseppe G.J.T., Schoenborn M.A., Armitage R.J.,
RA Fanslow W.C., Nakajima T., Baker E., Sutherland G.R., Poindexter K.,
RA Birks C., Albert A., Friend D., Gimpel S.D., Gayle R.B. III;
RT "The CD39 lymphoid cell activation antigen. Molecular cloning and
RT structural characterization."
RL J. Immunol. 153:3574-3583(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SvJ;
RX MEDLINE=98399871; PubMed=9730622;
RA Schoenborn M.A., Jenkins N.A., Copeland N.G., Gilbert D.J.,
RA Gayle R.B. III, Maliszewski C.R.;
RT "gene structure and chromosome location of mouse Cd39 coding for an
RT ecto-apyrase."
RL Cytogenet. Cell Genet. 81:287-289(1998).
CC -1- FUNCTION: In the nervous system, could hydrolyze ATP and other
CC nucleotides to regulate purinergic neurotransmission. Could also
CC be implicated in the prevention of platelet aggregation.
CC Hydrolyzes ATP and ADP equally well.
CC -1- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 phosphate.
CC -1- COFACTOR: Requires calcium and magnesium (By similarity).
CC -1- SUBUNIT: Homodimer; disulfide-linked (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: Belongs to the GDA1 / CD39 NTPase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC -----
CC EMBL; AF037366; AAB92259.1; -
CC EMBL; AF041818; AAC83203.1; -
CC EMBL; AF041812; AAC83203.1; JOINED.
CC EMBL; AF041813; AAC83203.1; JOINED.
CC EMBL; AF041814; AAC83203.1; JOINED.

```

```

DR EMBL; AF041815; AAC83203.1; JOINED.
DR EMBL; AF041816; AAC83203.1; JOINED.
DR EMBL; AF041817; AAC83203.1; JOINED.
DR MGD; MGI:102805; Entpdl
DR GO; GO:0005605; C:basal lamina; IDA.
DR GO; GO:0004050; P:pyruvate activity; IDA.
DR GO; GO:0006206; P:ATP catabolism; IDA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IDA.
DR GO; GO:0030168; P:platelet activation; IDA.
DR GO; GO:0009181; P:purine ribonucleoside diphosphate catabolism; IDA.
DR InterPro: IPR000407; GDA1_CD39_NTPase.
DR Pfam: PF01150; GDA1_CD39_1.
DR PROSITE; PS01238; GDA1_CD39_NTPase; 1.
KW Hydrolase; Transmembrane; Antigen; Glycoprotein; Calcium; Magnesium.
FT DOMAIN 1 16 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 17 37 POTENTIAL.
FT DOMAIN 38 478 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 479 499 POTENTIAL.
FT DOMAIN 500 510 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 73 73 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 226 226 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 291 291 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 428 428 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 457 457 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 510 AA; 57205 MW; 8E6A6113D2E13930 CRC64;

Query Match 16.4%; Score 370; DB 1; Length 510;
Best Local Similarity 27.1%; Pred. No. 7.3e-22;
Matches 121; Conservative 71; Mismatches 173; Indels 82; Gaps 19;

QY 32 GTFLSMCPINVSASTLYGIMFDAGSTGTRIHVTVFVKMPQQLPILEGEVDSVK-PGL 90
Db 36 GUTQKPLPENVK-----YQIVLDAGSSHTNLYIYKPAEKENDTGVVQQLSECVKGPFI 91
QY 91 SAFVDQPKQAGTVOGLLEVAKDSIPRSHWKTTPVVKATAGLRLL-----PEHAKALLPE 147
Db 92 SKVAQKTDEIGAYLAECMELSTELIPTSKHQTPVYLGATAGMLRLMESEQSADEVLA 151
QY 148 YKEIFRKPPLVPKGSVIMDSDEGILAWTVNLTGOL-----HGRQETVG 196
Db 152 VSTSLKSYPP--DFQAKIITQBEQAYGWIITNLLGRFTQEQSWLSLISDSQKETPG 209
QY 197 TLDLAGASTQITFLPQFQKTLFOTPRGYLTSEMFENSTYKLYTHSYLFGGLKAARLATIG 256
Db 210 ALDLGASTQITFVPQ-NSTIE-SPENSL-QFLYGEDYTVYTHSFCLYCKDQALWQKL- 265
QY 257 ALETSGTCHTFRSACL-----PRWLEAWIFGVKQYQVGNQGEV 298
Db 266 AKDIQVSSGGVILKPCFNPGEKVVNVSELYTPTCTKFEKLPDQFRIQTGD----- 320
QY 299 GPEPCVAEVLRVVRGKLHQP-E-----VORGFSFYAFSYTYDRAVTDMDY----- 344
Db 321 -VEQCHQSILELFNNS-HCPYSQCAFNGVFLPPLHGSFGAFSAFY-----PVMDFKKV 372
QY 345 EKGKILKVEDFERKAREVCD-NLENTSGSPPL-----CMDLSVITALLKDGFGFA 395
Db 373 AKNSVISQKMTETITKFNFCSSWEETKTSYPSVKEKYLSEYCFSGAYILSLQ-GYNFTD 431
QY 396 STVLQI--TKKVNNIETGWALGATPHL 420
Db 432 SWEQHPMGKIDSNAGWTLYMLNL 458

RESULT 14
ENPI_BOVIN STANDARD; PRT; 513 AA.
AC O18956;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 1 (EC 3.6.1.5)
DE (NTPDase1) (Ecto-ATP diphosphohydrolase) (ATPDase) (lymphoid cell

```

DE activation antigen) (CD39 antigen) (Ecto-apyrase).
GN ENTED1 OR CD39.
OS Bos taurus (Bovine).
OC Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Aortic endothelium;
RA Chang A.S., Garcia R.L., Chang S.M., Schilling W.P.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 97-103; 123-133; 136-140; 145-168 AND 459-471.
RC TISSUE=Aorta;
RX MEDLINE=97115958; PubMed=8955160;
RA Kaczmarek E., Kozlak K., Seivigny J., Siegel J.B., Anrather J.,
RA Beaudoin A.R., Bach F.H., Robson S.C.;
RT "Identification and characterization of CD39/vascular ATP
RT diphosphohydrolase."
RL J. Biol. Chem. 271:33116-33122(1996).
CC -!- FUNCTION: In the nervous system, could hydrolyze ATP and other
CC nucleotides to regulate purinergic neurotransmission. Could also
CC be implicated in the prevention of platelet aggregation.
CC Hydrolyzes ATP and ADP equally well.
CC -!- CATALYTIC ACTIVITY: ATP + 2 H₂O = AMP + 2 phosphate.
CC -!- COFACTOR: Requires calcium and magnesium.
CC -!- SUBUNIT: Homodimer; disulfide-linked (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: Belongs to the GDAL / CD39 NTPase family.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; AF005940; AAB62382.1; -
DR InterPro; IPR000407; GDAL CD39_NTPase.
DR Pfam; PF01150; GDAL CD39_1.
DR PROSITE; PS01238; GDAL CD39_NTPASE; 1.
KW Hydrolase; Transmembrane; Antigen; Glycoprotein; Calcium; Magnesium.
FT DOMAIN 1 16 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 17 37 POTENTIAL.
FT DOMAIN 38 481 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 482 502 POTENTIAL.
FT DOMAIN 503 513 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 227 227 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 245 245 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 373 373 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 460 460 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 97 97 K -> N (IN REF. 2).
FT CONFLICT 101 103 INV -> CGF (IN REF. 2).
FT CONFLICT 464 484 K -> V (IN REF. 2).
SQ SEQUENCE 513 AA; 58113 MW; 20FE98F27B6D3F96 CRC64;
Query Match 16.2%; Score 365; DB 1; Length 513;
Best Local Similarity 26.5%; Pred. No. 1.9e-21;
Matches 129; Conservative 70; Mismatches 180; Indels 108; Gaps 19;
QY 8 VTFMLVVCVCAVSHRNQOTWFEGIFLSSMCPINVSASTLYGIMFDAGSTTRHVIYTF 67
DB 17 ILSLGFSCIIAIVIA-----LLALGLTKQKALPENVK-----FGVLDAGSSSTSYIYRW 67
QY 68 VOKMPCQLPILEGEVDSVK-PGLSAFVDPQKQAGTVOGLLEVAKDSIPRSHWKXTPVY 126
DB 68 PREKENDTGWVQIESNVKVGISGFACKVNEINVILTACHERAQKVIPIQHWETPY 127

QY 127 LKATAGRLLL-PEHK--AKALLPEVKIIPKSPFLVPKGSVIMDGSDEGILAWTVNVL 183
DB 128 LGATAGRLLLRMENKQADKILAAVASSISSEYFP--DFQARIISGGEAGYHITVNYL 185
QY 184 TQQL-----HGRQETVGTLDGSASTQITFLPQFEKTLSEQTTPRGYITSFPMF 231
DB 186 LKQFQKLSWFLNFKSPKDDTQETYGALDGGASTQITFVPQNETT--ESPNNL-YFRLY 242
QY 232 NSTYKLYTHSVLGFGLXAARLATLGALETETDQ----- 265
DB 243 GKNSVYTHSFLCYGKQDALLQKL-ALGLOQTNGIIEHPCPHSYMRKIKMSVLNEGFT 301
QY 266 --HTRFSACLPRLWEAEWIFGCVKYQYCGNCGEVGPEPCYAEVLRV-----V 311
DB 302 KRHELNSFYB-----LVDIEIRGAGN-----FQRCFQSIQLFNTSCPVSSCSF 347
QY 312 RKLHQPEEVGSGFYAFSYIYDRAVTDMDIDYKGGILKVEDFERKAREVC----- 363
DB 348 NGVFLPPLHGGQFGAFSAFYV-----MEFLNLTSSEBSVVEQTEKLEEFCAQRWEVQ 401
QY 364 ----DNLENFTSGSPFLCMDSYITALLKQDGF-ADS-TVLQLTKKVNNIETGALGAT 417
DB 402 KNFGEVKEKYLSE---YCFSGTYILVLLNGYHFTASSKXNIHFMKVRSIDVGWTLGYM 458
QY 418 PHLLQSL 424
DB 459 LNLTKNI 465
RESULT 15
ENP4_MOUSE
ID ENP4_MOUSE STANDARD; PRT; 613 AA.
AC Q9DBT4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 4 (EC 3.6.1.6)
DE (NTPDase4) (Uridine-diphosphatase) (UDPase) (lysosomal apyrase-like
DE protein of 70 kDa).
GN LYSAL1 OR ENTPD4 OR LALP70.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yananaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann M., Gaasterland T., Glissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boiffelli D., Bojunga N., Carninci P., de Bonaldo M.P.,
RA Brownstein M.J., Burt C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustinchich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombauts P.,
RA Nordone P., Ring B., Schoenwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Yashaw-Baris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT *functional annotation of a full-length mouse cDNA collection*;
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Brain, and Mammary gland;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong E.,
RA Statchenko L., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RL human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
RX MEDLINE=20317099; PubMed=10858452;
RA Biederick A., Kusan C., Kunz J., Blaeser H.-P.;
RT "First asparase splice variants have different enzymatic properties.";
RL J. Biol. Chem. 275:19018-19024(2000).
CC -!- FUNCTION: Hydrolyzes preferentially nucleoside 5'-diphosphates,
CC nucleoside 5'-triphosphates are hydrolyzed only to a minor extent
CC (by similarity).
CC -!- CATALYTIC ACTIVITY: A nucleoside diphosphate + H(2)O = a
CC nucleotide + phosphate.
CC -!- COFACTOR: Requires calcium and magnesium.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC Localizes in the Golgi and autophagic vacuoles/lysosomes (By
CC similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1; Synonyms=LALP70;
CC IsoId=Q9DBT4-1; Sequence=Displayed;
CC Name=2; Synonyms=LALP70V;
CC IsoId=Q9DBT4-2; Sequence=VSP_003615;
CC -!- TISSUE SPECIFICITY: Ubiquitous.
CC -!- SIMILARITY: Belongs to the GDAL / CD39 NTPase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AK004761; BAB23542.1; -;
CC EMBL; BC006924; AAB06924.1; -;
CC EMBL; BC043134; AAB43134.1; -;
CC MGD; MGI:1914714; Lysal1.
CC InterPro; IPR000407; GDAL_CD39_NTPase.
CC Pfam; PF01150; GDAL_CD39_1.
CC PROSITE; PS01238; GDAL_CD39_NTPASE; FALSE NEG.
CC Hydrolase; Transmembrane; Glycoprotein; Calcium; Magnesium;
CC Alternative splicing; Golgi stack; Lysosome.
CC DOMAIN 1 33 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 34 54 POTENTIAL.
CC DOMAIN 55 559 LUMENAL (POTENTIAL).
CC TRANSMEM 560 580 POTENTIAL.
CC DOMAIN 581 613 CYTOPLASMIC (POTENTIAL).
CC CARBOHYD 404 404 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 407 407 N-LINKED (GLCNAC. .) (POTENTIAL).
CC VARSPLIC 287 294 Missing (in isoform 2).
CC /FTID=VSP_003615.
CC SEQUENCE 613 AA; 69745 MW; DES28F512ABEF52P CRC64;
SQ

Query Match 16.0%; Score 360.5; DB 1; Length 613;
Best Local Similarity 24.4%; Pred. No. 5.4e-21;

Matches 127; Conservative 77; Mismatches 185; Indels 131; Gaps 21;
QY 11 MLVWCVCVAVS-----ERNQOTWFE-----GIFLSSMCPINV-----SASTLYGIMF 53
DB 34 IIVISILAAASLLYFVSVIIIRSKYGLSKDKKQFYRLARVTDVBATNNPVSVMYIVV 93
QY 54 DAGSTGTIRHVYTFVQKMPGQLPIL-----GEVDSVKPGLSAFVDQPKQAETV 104
DB 94 DCGSSGSRIFVYCWPHNGNPHDLDIRQMDKQRKPVVMKIKPGISEFATSPKVSVDYI 153
QY 105 QGLLIVAKDSIPRSHWKTTPVVLKATAGLRLLPEHKALLFEVKEIFRKSP-----FLVP 160
DB 154 SPFLSFAAHVPRAKHETPLYILCTAGMRVLPSQKAIL-----BDLLTDIPVHYDFLFS 210
QY 161 KGSVIMDGSDEGILAWTVVAFLTGQLHGH-----RQRTVGT 197
DB 211 DSHAEVIGKQGVVAMTGINFVLRPF-HIEEDAEVVEVNIPOSESEALVRKTAGV 269
QY 198 LDLGASTQITF-LPQ-----FEKTLQTPRGYLSFEM-----FNSYKLYTHSYLQFG 246
DB 270 LDMGGVSTQIAYEVPQTVSFASSQOEBAKXLLAEFNLCGVHQTGHEVYRVVYVATFLQFG 329
QY 247 LKAARSLA-----TLGALETSGTGHFTFRSACLPRWLEAWIFGQVKYQYGN 293
DB 330 GNAARQRYEDRLFASTVQKNRLLGKQTGLTPDAPLLDCLPLDIKDE-----LQNGQ 382
QY 294 Q---EGEVGFPCYAEVLRVVRGKLHQ-----PEEVQSGSPYAFSYXDRAVDT 339
DB 383 TLYLQGTGDFDLK-RETLQPFANKTNETQTSINGVYQPPINFQSEFYGFSEFYCTEDV 441
QY 340 DMIDYEKGILKVEDFERKAEVCDN-----LENPTSG-----SPFLCMDLSYIT 384
DB 442 ----LRMGDVNAARFTQAADKYCATKWSILRRPDRGLYASHADLHLKLYQCFKSAWVF 497
QY 385 ALLKDGFGFADS-----TVLQLTQKVNNIETGWALGATFH 419
DB 498 EVFHKGFSPFVTKMLKALQVYDK-----EVQWTLGALLY 533

Search completed: July 1, 2004, 13:47:09
Job time : 12 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2004, 13:43:06 ; Search time 38.5 Seconds
(without alignments)
3507.579 Million cell updates/sec

Title: US-10-091-085-3
Perfect score: 2250
Sequence: 1 MATSGTTFVFMVVCVCSA.....ETGVALGATFLLQSLGISH 428

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL_25.*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mmc.*
 - 8: sp_organelle.*
 - 9: sp_phage.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_rvirus.*
 - 16: sp_bacteriap.*
 - 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2236	99.4	428	Q96RX0	Q96rx0 homo sapien
2	2104	93.5	407	Q8WUB3	Q8wub3 homo sapien
3	1990.5	88.5	427	Q8CD29	Q8cd29 mus musculus
4	1984.5	88.2	427	Q8BR23	Q8br23 mus musculus
5	996	44.3	483	Q8TAS7	Q8tas7 homo sapien
6	996	44.3	484	Q725B5	Q725b5 homo sapien
7	996	44.3	503	Q8N3H3	Q8n3h3 homo sapien
8	696.5	31.0	461	O76268	O76268 drosophila
9	696.5	31.0	464	O9VQI8	O9vqi8 drosophila
10	616.5	27.4	479	O9XU84	O9xu84 caenorhabdi
11	593	26.4	278	Q8CH23	Q8ch23 mus musculus
12	538.5	23.9	556	Q9UT35	Q9ut35 schizosacch
13	518.5	23.0	455	Q84UE0	Q84ue0 medicago tr
14	518.5	23.0	489	Q8H7L6	Q8h7l6 oryza sativ
15	515.5	22.9	599	Q8TGH6	Q8tgh6 candida alb
16	511	22.7	522	Q9HEM6	Q9hem6 kluyveromy

17	507.5	22.6	455	10	Q9SPM6	Q9spm6 medicago sa
18	503.5	22.4	467	10	Q9SPM7	Q9spm7 dolichos bi
19	498	22.1	462	10	Q9XFC9	Q9xlc9 dolichos bi
20	498	22.1	466	10	Q84UE2	Q84ue2 medicago tr
21	492	21.9	467	10	Q84UD8	Q84ud8 medicago tr
22	491.5	21.8	456	10	Q9SPM8	Q9spm8 lotus japon
23	491	21.8	455	10	Q84L88	Q84l88 glycine max
24	489	21.7	455	10	Q9PEA6	Q9pea6 pisum sativ
25	489	21.7	472	10	Q8L704	Q8l704 arabidopsis
26	488	21.7	463	10	Q9FVC3	Q9fvc3 glycine soj
27	488	21.7	472	10	Q9SPM5	Q9spm5 arabidopsis
28	486	21.6	455	10	Q8GTB1	Q8gtb1 pisum sativ
29	485	21.6	472	10	Q9M7B3	Q9m7b3 arabidopsis
30	484	21.5	471	10	Q9SQG2	Q9sqg2 arabidopsis
31	481	21.4	447	10	Q9SLV4	Q9slv4 pisum sativ
32	480.5	21.4	467	10	Q84UE1	Q84ue1 medicago tr
33	476.5	21.2	455	10	Q84UD9	Q84ud9 medicago tr
34	475	21.1	468	10	Q9FVC2	Q9fvc2 glycine soj
35	473	21.0	467	10	Q8RTV6	Q8rtv6 pisum sativ
36	472	21.0	407	10	Q9AVN8	Q9avn8 pisum sativ
37	472	21.0	455	10	Q8RVU0	Q8rvu0 pisum sativ
38	465	20.7	447	10	Q9FUI1	Q9fui1 pisum sativ
39	464.5	20.6	558	3	Q8TGS8	Q8tgs8 aspergillus
40	462	20.5	466	10	Q9AUI5	Q9aul5 medicago tr
41	454	20.2	454	10	Q84UE3	Q84ue3 medicago tr
42	449	20.0	455	10	Q8RTV9	Q8rtv9 pisum sativ
43	443	19.7	473	10	Q8RTV8	Q8rtv8 pisum sativ
44	436.5	19.4	556	3	Q9C2M0	Q9c2m0 neurospora
45	435.5	19.4	271	5	Q8IPZ6	Q8ipz6 drosophila

ALIGNMENTS

RESULT 1

Q96RX0 PRELIMINARY; PRT; 428 AA.

AC Q96RX0;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Pcp proto-oncogene protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MSBLINE=20173601; PubMed=10708485;
RA Recio J.A., Zambrano N., Pena, Ld, Reig J.A., Rhoads A., Rouzaut A.,
RA Notario V.;
RT "The human PCP proto-oncogene: cDNA identification, primary
RT structure, chromosomal mapping, and expression in normal and tumor
RT cells.";
RL Mol. Carcinog. 27:229-236(2000).
DR EMBL; AF136572; AAR2950.1; "
DR InterPro; IPR000407; GDAI_CD39_NTPase.
DR Pfam; PF01150; GDA_CD39; 1.
SQ SEQUENCE 428 AA; 47431 MW; P2C4F7DE650A44F6 CRC64;

Query Match 99.4%; Score 2236; DB 4; Length 428;
Best Local Similarity 99.5%; Pred. No. 6.4e-184;
Matches 426; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	1	MATSGTTFVFMVVCVCSA	VSHNQQTWFGI	FLSSMCPINV	SASTLYGIMFDAGSTGT	60
Db	1	MATSGTTFVFMVVCVCSA	VSHNQQTWFGI	FLSSMCPINV	SASTLYGIMFDAGSTGT	60
Qy	61	RIHYVTFVQKMPGOLPI	EGEVFDSVKPGLSAF	VDQKQGAETVQGL	LEVAKDSIPRSHW	120
Db	61	RIHYVTFVQKMPGOLPI	EGEVFDSVKPGLSAF	VDQKQGAETVQGL	LEVAKDSIPRSHW	120
Qy	121	KKTPTVVKATAGLRLLP	EHKAKALLFVKEIFR	SPFLVPKGSVIMD	GSDEGILAWTV	180

```
Db 121 KKTPTVVKATAGLRLLPEHKAALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180
Qy 181 NFLTGQHGHRQSTVGTGLDGGASTQITFLPQPEKTLQTPRGYLTSEFNFSTYKLYTH 240
Db 181 NFLTGQHGHRQSTVGTGLDGGASTQITFLPQPEKTLQTPRGYLTSEFNFSTYKLYTH 240
Qy 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQGEVGF 300
Db 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQGEVGF 300
Qy 301 EPCYAEVLVRGKLGHPQEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 301 EPCYAEVLVRGKLGHPQEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
Qy 361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400
Db 361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400
Qy 421 LQSLGISH 428
Db 421 LQSLGISH 428

RESULT 2
Q8WUB3 PRELIMINARY; PRT; 407 AA.
AC Q8WUB3;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Similar to ectonucleoside triphosphate diphosphohydrolase 5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC020966; AAH20966.1;
DR GO; GO:0016787; P:hydrolase activity; IEA.
DR InterPro; IPR000407; GDAI_CD39_NTPase.
DR Pfam; PFO1150; GDAI_CD39_1.
KW Hydrolase.
SQ SEQUENCE 407 AA; 45336 MW; D92A5F7DC9EC9E5B CRC64;

Query Match 93.5%; Score 2104; DB 4; Length 407;
Best Local Similarity 100.0%; Pred. No. 1.4e-172;
Matches 400; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATSWGTVFFMLVWSCVCSAVSHRNQQTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSWGTVFFMLVWSCVCSAVSHRNQQTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Qy 61 RIHVYTFVQKMPGQPLILEGEVDSVKPGLSAFVDQPKQAEIVQGLLEVAKDSIPRSHW 120
Db 61 RIHVYTFVQKMPGQPLILEGEVDSVKPGLSAFVDQPKQAEIVQGLLEVAKDSIPRSHW 120
Qy 121 KKTPTVVKATAGLRLLPEHKAALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180
Db 121 KKTPTVVKATAGLRLLPEHKAALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180
Qy 181 NFLTGQHGHRQSTVGTGLDGGASTQITFLPQPEKTLQTPRGYLTSEFNFSTYKLYTH 240
Db 181 NFLTGQHGHRQSTVGTGLDGGASTQITFLPQPEKTLQTPRGYLTSEFNFSTYKLYTH 240
Qy 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQGEVGF 300
Db 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQGEVGF 300
Qy 301 EPCYAEVLVRGKLGHPQEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 301 EPCYAEVLVRGKLGHPQEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
Qy 361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400
Db 361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400
Qy 421 LQSLGISH 428
Db 421 LQSLGISH 428

RESULT 3
Q8CD29 PRELIMINARY; PRT; 427 AA.
AC Q8CD29;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 5.
GN ENTPD5
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK031581; BAC27461.1;
DR MGD; MGI:1321385; Entpd5.
DR InterPro; IPR000407; GDAI_CD39_NTPase.
DR Pfam; PFO1150; GDAI_CD39_1.
SQ SEQUENCE 427 AA; 47101 MW; 6E3773C842B58477 CRC64;

Query Match 88.5%; Score 1990.5; DB 11; Length 427;
Best Local Similarity 88.1%; Pred. No. 9.1e-163;
Matches 376; Conservative 25; Mismatches 25; Indels 1; Gaps 1;

Qy 1 MATSWGTVFFMLVWSCVCSAVSHRNQQTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSWGTVFFMLVWSCVCSAVSHRNQQTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Qy 61 RIHVYTFVQKMPGQPLILEGEVDSVKPGLSAFVDQPKQAEIVQGLLEVAKDSIPRSHW 120
Db 61 RIHVYTFVQKMPGQPLILEGEVDSVKPGLSAFVDQPKQAEIVQGLLEVAKDSIPRSHW 120
Qy 121 KKTPTVVKATAGLRLLPEHKAALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180
Db 121 KKTPTVVKATAGLRLLPEHKAALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180
Qy 181 NFLTGQHGHRQSTVGTGLDGGASTQITFLPQPEKTLQTPRGYLTSEFNFSTYKLYTH 240
Db 181 NFLTGQHGHRQSTVGTGLDGGASTQITFLPQPEKTLQTPRGYLTSEFNFSTYKLYTH 240
Qy 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQGEVGF 300
Db 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQGEVGF 300
Qy 301 EPCYAEVLVRGKLGHPQEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 301 EPCYAEVLVRGKLGHPQEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
Qy 361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400
Db 361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400
Qy 421 LQSLGISH 427
Db 421 LQSLGISH 426
```

```
Db 301 EPCYAEVLVRGKLGHPQEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
Qy 361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400
Db 361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400

RESULT 3
Q8CD29 PRELIMINARY; PRT; 427 AA.
AC Q8CD29;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 5.
GN ENTPD5
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK031581; BAC27461.1;
DR MGD; MGI:1321385; Entpd5.
DR InterPro; IPR000407; GDAI_CD39_NTPase.
DR Pfam; PFO1150; GDAI_CD39_1.
SQ SEQUENCE 427 AA; 47101 MW; 6E3773C842B58477 CRC64;

Query Match 88.5%; Score 1990.5; DB 11; Length 427;
Best Local Similarity 88.1%; Pred. No. 9.1e-163;
Matches 376; Conservative 25; Mismatches 25; Indels 1; Gaps 1;

Qy 1 MATSWGTVFFMLVWSCVCSAVSHRNQQTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSWGTVFFMLVWSCVCSAVSHRNQQTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Qy 61 RIHVYTFVQKMPGQPLILEGEVDSVKPGLSAFVDQPKQAEIVQGLLEVAKDSIPRSHW 120
Db 61 RIHVYTFVQKMPGQPLILEGEVDSVKPGLSAFVDQPKQAEIVQGLLEVAKDSIPRSHW 120
Qy 121 KKTPTVVKATAGLRLLPEHKAALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180
Db 121 KKTPTVVKATAGLRLLPEHKAALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180
Qy 181 NFLTGQHGHRQSTVGTGLDGGASTQITFLPQPEKTLQTPRGYLTSEFNFSTYKLYTH 240
Db 181 NFLTGQHGHRQSTVGTGLDGGASTQITFLPQPEKTLQTPRGYLTSEFNFSTYKLYTH 240
Qy 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQGEVGF 300
Db 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQGEVGF 300
Qy 301 EPCYAEVLVRGKLGHPQEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 301 EPCYAEVLVRGKLGHPQEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
Qy 361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400
Db 361 EVCDNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400
Qy 421 LQSLGISH 427
Db 421 LQSLGISH 426
```

```
RESULT 4
Q8BR23 ID Q8BR23 PRELIMINARY; PRT; 427 AA.
AC Q8BR23;
DT 01-MAR-2003 (TEMBLrel. 23, Created)
DT 01-MAR-2003 (TEMBLrel. 23, last sequence update)
DT 01-JUN-2003 (TEMBLrel. 24, last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 5.
GN ENTPD5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Brain;
RC MDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK045828; BAC32507.1; -.
DR MGD; MGI:1321385; Entpd5.
DR InterPro; IPR000407; GDA1_CD39_NTPase.
DR Pfam; PF01150; GDA1_CD39_1.
DR SEQUENCE 427 AA; 47123 MW; 87BF2CC1CC1FCB9 CRC64;
SQ
Query Match 38.2%; Score 1984.5; DB 11; Length 427;
Best Local Similarity 87.8%; Pred. No. 3e-162;
Matches 375; Conservative 24; Mismatches 27; Indels 1; Gaps 1;
QY 1 MATSWGVFVPLVSVCSAVSRNQOTWEPGLFLSMCPINVSASTLYGIMFDAGSTGT 60
DB 1 MATSWGAV-FPLLIACVGSIVFREQOTWEPGLFLSMCPINVSAGTFYGIMFDAGSTGT 59
QY 61 RIHVYTFVQKMPQLPILEGEVDSVKPGLSAFVDPKQGAETVQGLLELVAKDSIPRSHW 120
DB 60 RIHVYTFVQKMPQLPILEGEVDSVKPGLSAFVDPKQGAETVQGLLELVAKDSIPRSEW 119
QY 121 KTFPVVLKATAGLRIPEHKAKALLFVKELFRKSPFLVPKGSVIMDSGDEGILAWTV 180
DB 120 ERTPEXLLKATAGLRLPEKAAQALLLEVEEIFKNSPFLVPDGSVIMDSGYEGILAWTV 179
QY 181 NPLTGLHGRQETVGTDLGGASTQITFLPQEKTLQTPRGYLSFEMFNSTFKLYTH 240
DB 180 NPLTGLHGRQETVGTDLGGASTQITFLPQEKTLQTPRGYLSFEMFNSTFKLYTH 239
QY 241 SYLGFGKKAARLATLGALETGDTGHTFRSACLPRMLEAEWIFGGVKYQYGGNQGEVGP 300
DB 240 SYLGFGKKAARLATLGALEAKGDTGHTFRSACLPRMLEAEWIFGGVKYQYGGNQGEVGP 299
QY 301 EPCVAEVLVRVGRKLPQEVORGSFVAFSYYYDRAVDMDIDYKGGILKVEDFERKAR 360
DB 300 EPCVAEVLVRVGRKLPQEVORGSFVAFSYYYDRAVDMDIDYKGGILKVEDFERKAR 359
QY 361 EVCDNLNFTSGSPFLCMLDSYITALLKDGFGFADSTVLQITKKVNNIETGVALGTFHL 420
DB 360 EVCDNLGSPSSGSPFLCMLDTYITALLKDGFGFADGTLTLQITKKVNNIETGVALGTFHL 419
QY 421 LQSLGIT 427
DB 420 LQSLGIT 426
RESULT 5
Q8TAS7 ID Q8TAS7 PRELIMINARY; PRT; 483 AA.
AC Q8TAS7;
DT 01-JUN-2002 (TEMBLrel. 21, Created)
DT 01-JUN-2002 (TEMBLrel. 21, last sequence update)
DT 01-JUN-2003 (TEMBLrel. 24, last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 6 (Putative
```

```
DE function).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RP TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC025980; AAH25980.1; -.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000407; GDA1_CD39_NTPase.
DR Pfam; PF01150; GDA1_CD39_1.
DR Hydrolase.
DR SEQUENCE 483 AA; 53119 MW; A850E5035BCDCE8F CRC64;
SQ
Query Match 44.3%; Score 996; DB 4; Length 483;
Best Local Similarity 52.2%; Pred. No. 5.4e-77;
Matches 203; Conservative 57; Mismatches 123; Indels 6; Gaps 4;
QY 40 PINVSA---STLYGIMFDAGSTGTRIHYTFVQKMPGQLPILEGEVDSVKPGLSAFVDQ 96
DB 89 PLGTAADGHEVYFYGIMFDAGSTGTRVHVVFQFT-RPRBTPPTLTHTFKALKPGLSAYADD 147
QY 97 PKQGAETVQGLLELVAKDSIPRSHWKTTPVLKATAGLRLPEHKAKALLFVKELFRKSP 156
DB 148 VEKSAQIGIRELLDVAKQDIPDFWKATPVLKATAGLRLPEKAKALLQKVKEVFKASP 207
QY 157 FLVPKGSVIMDSGDEGILAWTVNPLTGLHGRQETVGTDLGGASTQITFLPQEKFT 216
DB 208 FLVGDGCVSINMGTDGVSAMITINFLTCSLKTPOGSSVGMGLDGGSGTQIAFLPRVEGT 267
QY 217 LBOTPRGYLTSPEMNSTVKLYTHSYLGFGKKAARLATLGALETE-GTDTGHTFRSACLPR 275
DB 268 LQASPPGLTALRMFNRYKLYSYLGFGKKAARLATLGGVEGQPAKDKELVSPCLSP 327
QY 276 MLEAEWIFGGVKYQYGGNQGEVGPFCVYAEVLVRVGRKLPQEVORGSFVAFSYYYDR 335
DB 328 SFKGEWEHAEVTVRVSGQKAAASLHELCAARVSEVLQNRVHRTVEEVKHYVDFVAFSYYYDL 387
QY 336 AVDTMDIDYKGGILKVEDFERKAREVCNLENFTSGSPFLCMLDSYITALLKDGFGFAD 395
DB 388 AAGVGLIDAKKGSVVGVDFAARKVCKTLESTQPSQSPSCMDLTVLSILLQE-RGFPR 446
QY 396 STVLQITKKVNNIETGVALGTFHLIQL 424
DB 447 SKVLKLRKIDNVETSWALGAIFHVIDS 475
RESULT 6
Q7Z5B5 ID Q7Z5B5 PRELIMINARY; PRT; 484 AA.
AC Q7Z5B5;
DT 01-OCT-2003 (TEMBLrel. 25, Created)
DT 01-OCT-2003 (TEMBLrel. 25, last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, last annotation update)
DE CD39L2 nucleotidase.
GN ENTPD6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RP TISSUE=Brain;
RA Ivanenkov V.V., Murphy-Piedmonte D.M., Kirley T.L.;
RT "Bacterial Expression, Characterization, and Disulfide Bond
RT Determination of Soluble Human NTPDase6 (CD39L2) Nucleotidase:
RT Implications for Structure and Function."
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY327581; AAP92131.1; -.
DR SEQUENCE 484 AA; 53274 MW; D00A5D2915DF36CE CRC64;
SQ
```

```
Query Match 44.3%; Score 996; DB 4; Length 484;
Best Local Similarity 52.2%; Pred. No. 5.4e-77;
Matches 203; Conservative 57; Mismatches 123; Indels 6; Gaps 4;

QY 40 PINVSA---STLYGIMFDAGSTGTRHYVTVQVMPGQPLILEGEVDSVKPGLSAFVDQ 96
DB 90 PLGTAADGHEVFYGINFMDAGSTGTRVHVTFQFT-RPPRETFTLTHTFFKALPKGLSAYADD 148
QY 97 PKQAGTVOGLLEVAKDSIPRSHWKTTPVVKATAGLRLLPEHKAALIFEVKEIPRKSP 156
DB 149 VKSAAGIRELLDVAQDIPDFWKATPLVKATAGLRLLPEKAKLQKVKYKVPKASP 208
QY 157 FLVPGKSVSMGDSGLIAWTVNFTLQGLHGRHOETVGTDLGASTQITFLPQPEKT 216
DB 209 FLVGDDCVSIMGDTDEGVSAMITNFTLGSKTGGSSVGVGLDGLGGSTQIAFLPRVEGT 268
QY 217 LEOTPRGYLTSPFENSTYKLYTHSYLGFGLKARLATIACALETE-CTDGHFRSACLPR 275
DB 269 LQASPPGYLTALRMFNRTYKLYSYLGLGMSARLAILGGVGPQAKDGKELVSPCLSP 328
QY 276 WLEAEWIFGVGYOYGNQGEVGPVPCVAEVLVRVGRKLGHPQEEVQVSGSFAPSYYYDR 335
DB 329 SFKGEWEHAEVTVRSQKAAASLHELCAARVSEVLQNRVHRTVEVKHVDFAFSYYDL 388
QY 336 AVDTMDIDYKGGILKVEDPERKAREVCDNLENFTSGSPFLDMDLSVITALLKDGFGPAD 395
DB 389 AAGVGLIDAEGKGLVGVGFIAAKYVCRTELETQPSPPSCMDLTVVSLILOE-FGPPR 447
QY 396 STVLQTLTKVNNIETGVALGATHLLQSL 424
DB 448 SKVLKLRKIDNVETSWALGAIPHIDSL 476

RESULT 7
Q8N3H3 PRELIMINARY; PRT; 503 AA.
AC Q8N3H3
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein (fragment).
GN DKF2P761J1915.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Amalgam;
RA Ansgorge W., Wirkner U., Mewes H.W., Weil B., Wiemann S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL834158; CAD38864.1;
DR InterPro: IPR00407; GDA1_CD39_NTPase.
DR Pfam: PF01150; GDA1_CD39; 1.
KW Hypothetical protein.
FT NON TER
SQ SEQUENCE 503 AA; 54763 MW; 1639333F9139D8D9P CRC64;

Query Match 44.3%; Score 996; DB 4; Length 503;
Best Local Similarity 52.2%; Pred. No. 5.7e-77;
Matches 203; Conservative 57; Mismatches 123; Indels 6; Gaps 4;

QY 40 PINVSA---STLYGIMFDAGSTGTRHYVTVQVMPGQPLILEGEVDSVKPGLSAFVDQ 96
DB 109 PLGTAADGHEVFYGINFMDAGSTGTRVHVTFQFT-RPPRETFTLTHTFFKALPKGLSAYADD 167
QY 97 PKQAGTVOGLLEVAKDSIPRSHWKTTPVVKATAGLRLLPEHKAALIFEVKEIPRKSP 156
DB 168 VKSAAGIRELLDVAQDIPDFWKATPLVKATAGLRLLPEKAKLQKVKYKVPKASP 227
QY 157 FLVPGKSVSMGDSGLIAWTVNFTLQGLHGRHOETVGTDLGASTQITFLPQPEKT 216
DB 157 FLVPGKSVSMGDSGLIAWTVNFTLQGLHGRHOETVGTDLGASTQITFLPQPEKT 216
```

```
DB 228 FLVGDDCVSIMGDTDEGVSAMITNFTLGSKTGGSSVGVGLDGLGGSTQIAFLPRVEGT 287
QY 217 LEOTPRGYLTSPFENSTYKLYTHSYLGFGLKARLATIACALETE-CTDGHFRSACLPR 275
DB 288 LQASPPGYLTALRMFNRTYKLYSYLGLGMSARLAILGGVGPQAKDGKELVSPCLSP 347
QY 276 WLEAEWIFGVGYOYGNQGEVGPVPCVAEVLVRVGRKLGHPQEEVQVSGSFAPSYYYDR 335
DB 348 SFKGEWEHAEVTVRSQKAAASLHELCAARVSEVLQNRVHRTVEVKHVDFAFSYYDL 407
QY 336 AVDTMDIDYKGGILKVEDPERKAREVCDNLENFTSGSPFLDMDLSVITALLKDGFGPAD 395
DB 408 AAGVGLIDAEGKGLVGVGFIAAKYVCRTELETQPSPPSCMDLTVVSLILOE-FGPPR 466
QY 396 STVLQTLTKVNNIETGVALGATHLLQSL 424
DB 467 SKVLKLRKIDNVETSWALGAIPHIDSL 495

RESULT 8
Q76268 PRELIMINARY; PRT; 461 AA.
AC Q76268
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NTPase protein (LD11641P).
GN NTPASE OR CG3059.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINS=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter B.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Efannkoch C., Baldwin D.,
RA Ballieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fessler C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glaser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang D., Lai Z.,
RA Kimble B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kianos I., Simpson M., Skupski M.P., Smith T.,
RA Spier B., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers B.W., Rubin G.M., Venter J.C.;
```

RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195 (2000).
RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE=98341119; PubMed=9676430;
RA Chadwick B.P., Frischauf A.-M.;
RT "The CD39-like gene family: identification of three new human members
RT (CD39L2, CD39L3, and CD39L4), their murine homologues, and a member of
RT the gene family from Drosophila melanogaster.";
RL Genomics 50:357-367 (1998).
RN (3)
RN SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayan A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarini H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuncio J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celnikier S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF041048; AAC39133.1; -;
DR EMBL: AF041048; AAC39133.1; -;
DR EMBL: AF041048; AAC39133.1; -;
DR Flybase; FBgn0024947; NTPase.
DR InterPro; IPR000407; GDAL_CD39_NTPase.
DR Pfam; PF01150; GDAL_CD39; 1.
SQ SEQUENCE 461 AA; 50845 MW; 27D00321F91A9DD1 CRC64;

Query Match 31.0%; Score 696.5; DB 5; Length 461;
Best Local Similarity 39.1%; Pred. No. 3.2e-51;
Matches 156; Conservative 74; Mismatches 122; Indels 47; Gaps 12;

QY 49 YGIMFDAGSTGTRIHVYTFVQKMPGQLPILGEVFDVSKPGLSAFVDPKQGAETVQGLL 108
Db 79 YAAIIDAGSTGSRVLAYKENRSFIDNKLVLVEELPKERKFGLSFPADNPAEGAHSLKLL 138

QY 109 EVAKDSIPRSHWKTTPVVLKATAGRLLEPHKAKALLFEVKETFRKSPFLVPKGSVIMD 168
Db 139 DEARAFIPKHSNSTPLVLTATAGRLLPASKAENILNAVRLFAKSEFSDVMDAVEIMD 198

QY 169 GSDEGILAWTVNFLTQGLHGRQETVGTLDLGASTQTTF-----LPQEKLEQTP 221
Db 199 GTDEGIFSWFTVNFLLGRLSKTNQ--AAALDLGGSGTQVTFSTPDQDPQVYVYKYMHEV- 255

QY 222 RGLTSEFMFNSTYKLYTHSYLGFGLKAARLATLGALETG--TDGHTFRSACL-PRWLE 278
Db 256 ---VTSSKKIN-----VFTHSYLGLGLMAARHAFV-----THGYKKEDTVLESVCVNPFIAN 304

QY 279 AEWIFGGVKYQYGGNQEGB-----VGPEPCYAEVLVRVVGKL-----HQPEVORGSP 326
Db 305 RTWTGNGVQYKVGKENGKSSAEQPIVDFDAC-----LELVKSKMPLVFKPKFTLKQHAV 360

QY 327 YAFSYTYDRAVDYDMDIDYKGGILKYVEDFERKAREVC--DNLENFTSGSPFLCMBLSYIT 384
Db 361 AAFSYTYFERAIESGLVDPLAGGETTVEAVRKAQEI CAIPNDE-----QPFMCFDLTFTS 415

QY 385 ALLKDGEPADSVTLQITKKNVNIETGMAWGATFHLLQS 423
Db 416 TLREGFGLMDGKKIKLYDKIDGHEISWALGCAINVLTS 454

RESULT 9

O9VQ18 PRELIMINARY; PRT; 464 AA.
AC O9VQ18;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE NTPASE protein.
GN NTPASE OR CG3059.
OS Drosophila melanogaster (fruit fly).
OC Sukarya; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Metapterygota; Dipera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayan A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Balles R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Kays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Svangelista C.C., Ferraz C., Ferreria S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Turner R., Venter E., Wang A.H., Wang X.,
RA Svitskas R., Tector C., Turner R., Venter E., Weissenbach J.,
RA Wang Z.-Y., Wasmann D.A., Weinstock G.M., Wu D., Yang S., Yao Q.A.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Zhao Q., Zheng L.,
RA Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhu X.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195 (2000).
DR EMBL: AB003581; AAF51182.1; -;
DR Flybase; FBgn0024947; NTPase.
DR InterPro; IPR000407; GDAL_CD39_NTPase.
DR Pfam; PF01150; GDAL_CD39; 1.
SQ SEQUENCE 464 AA; 51119 MW; 71D057AB85AE613D CRC64;

Query Match 31.0%; Score 696.5; DB 5; Length 464;
Best Local Similarity 39.1%; Pred. No. 3.2e-51;
Matches 156; Conservative 74; Mismatches 122; Indels 47; Gaps 12;

QY 49 YGIMFDAGSTGTRIHVYTFVQKMPGQLPILGEVFDVSKPGLSAFVDPKQGAETVQGLL 108
Db 82 YAAIIDAGSTGSRVLAYKENRSFIDNKLVLVEELPKERKFGLSFPADNPAEGAHSLKLL 141

QY 109 EVAKDSIPRSHWKTTPVVLKATAGRLLEPHKAKALLFEVKETFRKSPFLVPKGSVIMD 168
Db 142 DEARAFIPKHSNSTPLVLTATAGRLLPASKAENILNAVRLFAKSEFSDVMDAVEIMD 201

QY 169 GSDEGILAWTVNFLTQGLHGRQETVGTLDLGASTQTTF-----LPQEKLEQTP 221
Db 202 GTDEGIFSWFTVNFLLGRLSKTNQ--AAALDLGGSGTQVTFSTPDQDPQVYVYKYMHEV- 258

QY 222 RGLTSEFMFNSTYKLYTHSYLGFGLKAARLATLGALETG--TDGHTFRSACL-PRWLE 278
Db 259 ---VTSSKKIN-----VFTHSYLGLGLMAARHAFV-----THGYKKEDTVLESVCVNPFIAN 307

QY 279 AEWIFGGVKYQYGGNQEGB-----VGPEPCYAEVLVRVVGKL-----HQPEVORGSP 326
Db 308 RTWTGNGVQYKVGKENGKSSAEQPIVDFDAC-----LELVKSKMPLVFKPKFTLKQHAV 363


```
Qy 327 YAFSYVYDRAVDTMDIYKGGILKVEDFERKAREVC--DNLENFTSGSPFLCMLDSYIT 384
Db 364 AAFSYFPERAIESGLVDPLAGGTTVEAYEKKAQETCAIPNDE-----QPFMCFDLTFIS 418
Qy 385 ALLKDGFGFADSTVLOLTKVNNIETGWCALGATPHLQOS 423
Db 419 TLLREGFLNDGKKIKYKIDGHEISWALGCAYNVLTS 457

RESULT 10
Q9XU84 PRELIMINARY; PRT; 479 AA.
ID AC Q9XU84
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE K08H10.4 protein.
GN K08H10.4
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Gardner A.E.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=99069613; PubMed=9851916;
RX none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z83113; CAB05544.1; -
DR PIR; T23508; T23508
DR Wormpep; K08H10.4; C018877.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002034; AIPM/Hcit synth.
DR InterPro; IPR000407; GDAL_CD39_NTPase.
DR Pfam; PF01150; GDAL_CD39; 1.
DR PROSITE; PS00815; AIPM_HOMOCIT SYNTH 1; 1.
SQ SEQUENCE 479 AA; 5384 MW; 7EDC02A9D54A48ED CRC64;

Query Match 27.4%; Score 616.5; DB 5; Length 479;
Best Local Similarity 34.3%; Pred. No. 2.7e-44;
Matches 150; Conservative 81; Mismatches 167; Indels 39; Gaps 16;

Qy 9 FFMVLS--CVCSAVSHRNQOTWTEGIFLSMCMFINVSASTLYGIMFDAGSTGRIHYVT 66
Db 6 FSIILLISFFSLLSVVTTKQY-WCHGDVNL---NOHTCRFFTVI:DAGSTGRLHLK 60
Qy 67 FVQR-----MPQLPILEGVDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSH 119
Db 61 FIHPDPAIASHGMFPK---VEKEIFQEVKPGLSFPAKSPSSAADSLEPFLQARKEVPHFM 117
Qy 120 WKTPVVLKATAGLRLPEHKAKALLPEVKE-IFRKSPPFLVPKGSVIMGSDGILAWV 178
Db 118 WEKTPITLKATAGLRLPGMADDLISVEERIFNSGFFAAPPDAVNVMPGSEGVYSWF 177
Qy 179 TVNFLTGLH-----GHR-----QSTVGTLDLGGASTQTITFLPQPKTELEOTPRGYLTSFE 229
Db 178 TINTILLETLFDTEVTGKHPAARHSVAADFLLGGSTQLTYPNNEAVFSEHV-GYERDID 236
Qy 230 MFNSTYKLYTHSYLGFGLKAARLATLGALETG---TQHTFPRACLPRLWE-AEWIFGVV 286
Db 237 FFGHILRFTSFGLNGLIARLML-QLETDNIEISTHQLITSCMPEGYQLTEWEY-AL 294
Qy 287 KYQYGGKQEGVEGPEPCVAEVLVVR-GKLHQPEEVQVGRSPYAFSYTYDRAVDTMDIYE 345
Db 295 KP-WNINGSSSHFSFSCVITKNFVSESEINHLRELKGPSVYLFESYFFDRLNLSGLVKN 353
Qy 346 KGGILKVEDFERKAREVC-----DNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQL 401
```

```
Db 354 EGGKILKROFFKEAABIAICRREKTEIDGSHWFWOCLDTIYSLLRGQYFEDNQPLV 413
Qy 402 TKKVNNIETGWCALGATP 418
Db 414 AKKIKGMEVSWGQGLAF 430

RESULT 11
Q8CHZ3 PRELIMINARY; PRT; 278 AA.
ID AC Q8CHZ3
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Similar to ectonucleoside triphosphate diphosphohydrolase 6.
GN ENTPD6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Mammary gland;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC038126; AAH38126.1; -
DR MGD; MGI:1202295; Entpd6.
DR GO; GO:0015787; F:hydrolase activity; IEA.
DR InterPro; IPR000407; GDAL_CD39_NTPase.
DR Pfam; PF01150; GDAL_CD39; 1.
DR Hydrolase.
RX SEQUENCE 278 AA; 30130 MW; 75A92DD1AC76297F CRC64;

Query Match 26.4%; Score 593; DB 11; Length 278;
Best Local Similarity 58.0%; Pred. No. 1.2e-42;
Matches 119; Conservative 31; Mismatches 49; Indels 6; Gaps 2;

Qy 49 YGIMFDAGSTGRIHYVTQVQKPGQLPILEGVDSVKPGLSAFVDQPKQGAETVQGLL 108
Db 74 YGIMFDAGSTGRIHYVQFA-RPPGETPLTETFKALPGLSAVADDVKEKSAQGIQELL 132
Qy 109 EVAKOSIPRSHKKTVPVLKATAGLRLPEHKAKALLPEVKEIFRKSPPFLVPKGSVIMD 168
Db 133 NVAQHIPPYDFWRKATPLVLKATAGLRLPEHKAKALLQKVKFPAKSPFLVGDDCVSIMN 192
Qy 169 GSDEGILAWTVNFLTGLHGRQETVGTLDLGGASTQTITFLPQPKTELEOTPRGYLTSF 228
Db 193 GTDEGVSAWITVNFLLGSLKTPGSSSVGMGLDGGSTQITFLPVEGTQLQASPPGHLTAL 252
Qy 229 EMFNSTYKLYTHSYLGFGLKAARLA 253
Db 253 QMFNRTYKLYSYRWV-----CSRLLA 272

RESULT 12
Q9UT35 PRELIMINARY; PRT; 556 AA.
ID AC Q9UT35
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative guanosine-diphosphatase (Guanosine diphosphatase).
GN SPAC324.08 OR GDPI.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=972b-;
RA Barrell B.G., Rajandream M.A., Quail M., Seegar K., Harris D.;
```


Db 84 STRXAVIFDAGSGSRVHVYCF-----DGNLDLLPIGKEIELEFKOKKPGLSAYAMDPQRAA 139
QY 102 ETVOQLLEVAKXDSIPRSHWKKTPVVLKATAGRLLEPEHKAKALLFEVKEIER-KSPFLVP 160
Db 140 KSLVSLLEAEKVPVLEBQTPVVGATAGRLGCTEKSEIQAQVNDLLQDKSFRSQ 199
QY 161 KGSVISMDGSEGLAWTVNPLTQQLHGHQRTVGTLDLGGASTQITFLPQFETLEQT 220
Db 200 PEWTVLDSGSEGAQFQWTTIYLLGNLQKPYSHTVGVVDLGGSGVQAYALS-EKDAGKA 258
QY 221 P-----RGYLTSPFEMFNSTYKLYTHSYLGRGLKAARLATI-GALETGTGHTTSSACLP 274
Db 259 PPVAGESEYKELLKGTITTYLVHVSRLRGLAARABIKAGE-----GNDYRN-CML 312
QY 275 RWLEARNIFGVKYQYQGNQGEVGFPCYAEVLVRVWVKLHQPE----- 319
Db 313 EGHGQVAYGDDIIFEASGLSSG-ASYSKRAVAVRAL--KVDEPACTHMKCTFGGVWNGG 369
QY 320 --EVORGSFYAFSYVYDRAVDTMDYKGGILKVE--DPERKAREVCD-NLENFTS---- 371
Db 370 GGDQKULFVA-SFFFDRAAGFVN-PRAPFAKXPSDFEEAARRVCKLVKXKDAQATYP 427
QY 372 ----GSPFLCMLSYITALLKDGFGFADSTVLQITKKV-----NNIETGVALGATFHLLQ 422
Db 428 DVSEENVFLCMLVYQVTLVLDGFGVDYQDITLVKKVPYSNSFVEAAMPFGSAIEVAS 487
QY 423 S 423
Db 488 S 488

RESULT 15

Q8TGH6 PRELIMINARY; PRT; 599 AA.
AC Q8TGH6;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Guanosine diphosphatase (EC 3.6.1.42).
GN GDA1.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RA Herrero A. B.
RT "The Golgi Gpase of the fungal pathogen Candida albicans Affects
RT morphogenesis, glycosylation and cell wall properties."
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ421721; CADI18870.1; -
DR GO; GO:0004382; P:guanosine diphosphatase activity; IEA.
DR GO; GO:0016787; P:hydrolase activity; IEA.
DR InterPro; IPR000407; GDA1_CD39_NTPase.
DR Pfam; PF01150; GDA1_CD39; 1.
DR PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
KW Hydrolase.
SQ SEQUENCE 599 AA; 65952 MW; D5BC3A4F6E1B646A CRC64;

Query Match 22.9%; Score 515.5; DB 3; Length 599;
Best Local Similarity 32.6%; Pred. No. 1.9e-35;
Matches 143; Conservative 65; Mismatches 148; Indels 83; Gaps 15;
QY 49 YGIMFDAGSTGTRIHVYTFVQMPGQPLILEGEVDFSVKPGLSAFVDPQKQAGATVQGLL 108
Db 163 YVWMIDAGSTGSRVHVYEFNTCV--KPPQLLSERFEMLKPGLSSTFDITVGAASLDPLL 220
QY 109 EVAXDSIPRSHWKKTPVVLKATAGRLLEPEHKAKALLFEVKEIERKS-PR-LVPKGSVSI 166
Db 221 EVALKVPKQKQSTPVAVKATAGRLLEPEHKAKALLFEVKEIERKS-PR-LVPKGSVSI 280
QY 167 MDGSEGLAWTVNPLTQQLHGHQRTVGTLDLGGASTQITFLPQFETLEQTTPRGYL 225

Db 281 MDGKDEGVYAWVTANYLLGNIGGKELPTAAVFDLGGGSTQIVFEPDYK--VDEVPPVDGE 338
QY 226 TS--FEMFNSTYKLYTHSYLGRGL-----KAARLATI-LGALETGTGHTTFR----- 269
Db 339 TKYHFTFGDNQYTLTYQPSHLGYGLMOGRNKNVQLVLKNKLSEINLQKYTEKVKGAKATV 398
QY 270 ---SACLPRWLEAE-----WIFGKVKYQYGG---NOEGEV 298
Db 399 DVSNPCIPPGVAKDVQVELGEDEFYVANKGPPSSKXSTVAGGSCRYLAERKVLNKDAEC 458
QY 299 GPPECYAEVLVRVWVKLHQPEEV-----QSGSFYAFSYVYDRAVDTMDYKGGILKVED 354
Db 459 TSKEPCSF-----NGVHOPSLRTRTNKNSDMTVFSFYDRTNPIGM-----PSSFSVEE 506
QY 355 FERKAREVC-----DNLENFTSGSPFLCMLSYITALLKDGFGFADSTVLQIT 402
Db 507 LKDLKLVCOGETFWKXIDLDHVKNLNE-EPQWCLDLSFITAMLTGVDYDIPHLRELKTA 565
QY 403 KKVNNIETGVALGATFHLL 421
Db 566 KTIDNNELGMCGLGASLPLL 584

Search completed: July 1, 2004, 13:48:39
Job time : 40.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2004, 13:41:36 ; Search time 52 Seconds
(without alignments)
2325.583 Million cell updates/sec

Title: US-10-091-085-3

Perfect score: 2250

Sequence: 1 MATSWGVPFPLVVCVCSA.....ETGWLGATPHLQLSLGISH 428

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

1: Geneseq1980s:*

2: Geneseq1990s:*

3: Geneseq2000s:*

4: Geneseq2001s:*

5: Geneseq2002s:*

6: Geneseq2003as:*

7: Geneseq2003bs:*

8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2250	100.0	428	3	AAV44849 Human CD3
2	2250	100.0	428	4	AAV72238 Human CD3
3	2250	100.0	428	4	AAV72243 Human CD3
4	2250	100.0	428	5	AAE19883 Human CD3
5	2235	99.3	428	3	AAV44850 Human CD3
6	2235	99.3	428	4	AAV72240 Human CD3
7	2104	93.5	405	3	AAV44851 Human CD3
8	2104	93.5	405	4	AAV72239 Human CD3
9	1832.5	81.4	465	5	AAE19884 Mouse CD3
10	1660	73.8	330	3	AAE19881 Human CD3
11	999	44.4	456	5	AAE19881 Human CD3
12	999	44.4	484	4	AAV72241 Human CD3
13	996	44.3	463	5	ABB06124 Human NS
14	996	44.3	467	5	ABJ04657 Protein o
15	992	44.1	456	4	AAV91929 Human p
16	989	44.0	446	5	ABJ04658 Protein o
17	941	41.8	450	7	ADC14220 Human enz
18	897	39.9	462	4	AAU30882 Novel hum
19	819.5	36.4	471	4	AAV72242 Mature hu
20	696.5	31.0	461	4	ABB66213 Drosophil
21	696.5	31.0	464	4	ABB59611 Drosophil
22	590	26.2	476	3	AAV70912 Human CD3
23	590	26.2	476	3	AAV70889 Protein e
24	533	23.7	476	3	AAV70911 Human CD3
25	533	23.7	476	3	AAV70888 Protein e

26	508.5	22.6	407	7	ABM73652 DNA clone
27	508.5	22.6	457	7	ABM74049 DNA clone
28	507.5	22.6	486	2	AAW85685 NBP46 roo
29	507.5	22.6	486	5	AAU78819 Lotus jap
30	499.5	22.2	467	2	AAW85687 DBX oligo
31	498	22.1	462	2	AAW85684 NBP46 roo
32	498	22.1	462	5	AAU78818 Dolichos
33	491.5	21.8	496	5	AAU78820 Medicago
34	489	21.7	455	4	AAE81952 Pea bligh
35	488	21.7	472	6	ABP81286 Arabidops
36	482.5	21.4	496	2	AAW85686 NBP46 roo
37	367	16.3	495	2	AAV33236 Human mem
38	358.5	15.9	458	4	AAU29271 Human PRO
39	358.5	15.9	458	4	ABU58647 Human PRO
40	358.5	15.9	458	6	ABU88195 Novel hum
41	358.5	15.9	458	6	ABU84510 Human sec
42	358.5	15.9	458	6	ABR66384 Human sec
43	358.5	15.9	458	6	ABR65774 Human sec
44	358.5	15.9	458	6	ABU99714 Human sec
45	358.5	15.9	458	6	ABU82953 Human PRO

ALIGNMENTS

RESULT 1

AAV44849

ID AAV44849 standard; protein; 428 AA.

XX

AC AAV44849;

XX

DT 18-MAY-2000 (first entry)

XX

DE Human CD39-L4 protein.

XX

CD39-L4; human; apyrase; nucleotide diphosphatase; NDase;

ATP diphosphohydrolase; ATPase; adenosine diphosphate; ADP; treatment;

platelet aggregation; antithrombotic; thrombosis; myocardial infarction;

cerebral ischemia; angina; vascular graft; extracorporeal circulation;

molecular weight marker; nutritional supplement; tumour; prevention;

drug targeting; Apyrase Conserved Region; ACR.

XX Homo sapiens.

XX

FH Key Location/Qualifiers

FT Peptide 1..22

FT Protein /label= Leader_peptide 23..428

FT /label= Mature_human_CD39-L4_protein

FT /note= Homologous to human and murine CD39"

FT Binding-site 54..58

FT Region /label= ATP_Binding_region 129..134

FT Region /label= Apyrase_Conserved_Region 169..173

FT Region /label= Apyrase_Conserved_Region 199..206

FT /note= "Conserved motif in ATPDases"

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PI Ford J, Mulero JJ;
XX WPI; 2000-192397/15.
DR N-PSDB; AAZ50356, AAZ50359.
XX New nucleic acid encoding human CD39-like protein, useful for treating
PT and preventing thrombotic disease.
XX
XX
PS Claim 15; Fig 2; 125pp; English.
XX
XX The present amino acid sequence is the CD39-L4 protein, an apyrase and/or
CC nucleotide diphosphatase (NDPase). It is isolated from the human foetal
CC liver-spleen cDNA library, B2HFUS20W. It is a soluble ATP
CC diphosphohydrolases (APDase) and is involved in the hydrolysis of
CC adenosine diphosphate (ADP), the agonist that causes platelet
CC aggregation. CD39-L4 protein has 30% and 80% homology to human and murine
CC CD39. It has platelet aggregation inhibition and antithrombotic activity.
CC CD39-L4 is used to treat or prevent thrombosis, myocardial infarction,
CC cerebral ischaemia and angina. It is also used in vitro, to maintain
CC vascular grafts or during extracorporeal circulation, to hydrolyse NDP,
CC as molecular weight markers and as nutritional supplements. It is used to
CC identify therapeutic agents that bind and modulate CD39-L4. It is coupled
CC to toxins for targeting drugs to tumours or other cells that express CD39
CC -L4
XX
XX Sequence 428 AA;

Query Match 100.0%; Score 2250; DB 3; Length 428;
Best Local Similarity 100.0%; Pred. No. 3.2e-208;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATSGTGVFFMLVSVCSAVSHRNQOTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
DB 1 MATSGTGVFFMLVSVCSAVSHRNQOTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
QY 61 RIHYVTFVQKMPGQLPILGEVFDSPVKGLSAFVDQPKQGAETVQGLLEKADSIIPRSHW 120
DB 61 RIHYVTFVQKMPGQLPILGEVFDSPVKGLSAFVDQPKQGAETVQGLLEKADSIIPRSHW 120
QY 121 KKTVPVLKATAGLRLPEHAKALLFEVKEIFRKSPTLVKPGSVSINDGSDGILAWTV 180
DB 121 KKTVPVLKATAGLRLPEHAKALLFEVKEIFRKSPTLVKPGSVSINDGSDGILAWTV 180
QY 181 NFLAGQLGHRQETVGTLDLGGASTQITFLPQFEKTLQTPRGYLTSEFNFNSTYKLYTH 240
DB 181 NFLAGQLGHRQETVGTLDLGGASTQITFLPQFEKTLQTPRGYLTSEFNFNSTYKLYTH 240
QY 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRMLEAEWIFGGVXYQYGGNQGEGVGF 300
DB 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRMLEAEWIFGGVXYQYGGNQGEGVGF 300
QY 301 BPCYAEVLVRVGRKLHQPVEVQSGSFYAFSYYYDRAVDVTMDIYKGGILKVEDFERKAR 360
DB 301 BPCYAEVLVRVGRKLHQPVEVQSGSFYAFSYYYDRAVDVTMDIYKGGILKVEDFERKAR 360
QY 361 EVCNDLENFTSGSPFLNDLSYITALLKDGFGFADSTVLQTKKNNIETGALGATFHL 420
DB 361 EVCNDLENFTSGSPFLNDLSYITALLKDGFGFADSTVLQTKKNNIETGALGATFHL 420
QY 421 LQSLGISH 428
DB 421 LQSLGISH 428

RESULT 2
AAB72238
ID AAB72238 standard; protein; 428 AA.
XX
AC AAB72238;
XX
DT 14-MAY-2001 (first entry)
XX
DE Human CD39 like protein CD39-L4 amino acid sequence.

XX Human CD39-like protein; apyrase; NDPase; platelet function inhibitor;
KW myocardial infarction; cerebral ischaemia; angina; arterial thrombosis;
KW cerebral artery thrombosis; platelet aggregation; inflammation;
KW apoptosis; autoimmune disorder; neurological disorder;
KW Alzheimer's disease; Parkinson's disease; cancer; CD39-L4.
XX
OS Homo sapiens.
XX
XX WO200110205-A1.
XX
PD 15-FEB-2001.
XX
XX 09-AUG-2000; 2000WO-US021790.
XX
XX 09-AUG-1999; 99US-00370265.
PR 11-JAN-2000; 2000US-00481238.
PR 25-APR-2000; 2000US-00557800.
PR 26-MAY-2000; 2000US-00583231.
PR 30-JUN-2000; 2000US-00608285.
XX
XX (HYSE-) HYSEQ INC.
XX
PI Ford J, Mulero JJ, Yeung G;
XX
XX WPI; 2001-147489/15.
DR N-PSDB; AAF63383.
XX
XX Polynucleotides encoding human CD39-like polypeptides, with apyrase
PT and/or NDPase activity, which are useful in the treatment of pathological
PT conditions caused by thrombosis (e.g. myocardial infarction) and
PT inflammatory disorders.
XX
XX Claim 15; Fig 2; 203pp; English.

This invention relates to polynucleotides encoding human CD39-like
polypeptides with apyrase and/or NDPase activity. The polypeptides having
APDase, including NDPase, activity are useful for inhibiting platelet
function and can therefore be used in the prophylaxis or treatment of
CC pathological conditions caused by or involving thrombosis or excessive
CC coagulation or excessive platelet aggregation, such as myocardial
CC infarction, cerebral ischaemia, angina, arterial thrombosis, cerebral
CC artery thrombosis or intracardiac thrombosis, and conditions associated
CC with venous thrombosis. CD39-L4 and CD39-L2 polypeptides are useful in
CC modulating disease states (including platelet aggregation, inflammation
CC and apoptosis) associated with ADP or other purinergic signalling by
CC reducing the levels of NDPs. The polypeptides are also useful for
CC prophylaxis or treatment of inflammation related disorders, such as
CC disorders involving sepsis or systemic inflammatory response syndrome or
CC SIRS (and associated conditions such as fever, tachycardia, tachypnea,
CC cytokine overstimulation); autoimmune disorders such as thrombosis,
CC atherosclerosis, acute pancreatitis, dermatitis, including psoriasis,
CC cirrhosis, reperfusion injury, asthma, multiple sclerosis, arthritis;
CC neurological disorders including neurodegenerative diseases, epilepsy,
CC depression, Alzheimer's disease, Parkinson's disease, Huntington's
CC disease, and amyotrophic lateral sclerosis; and cancer. The present
CC sequence represents human CD39 like protein CD39-L4

Query Match 100.0%; Score 2250; DB 4; Length 428;
Best Local Similarity 100.0%; Pred. No. 3.2e-208;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATSGTGVFFMLVSVCSAVSHRNQOTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
DB 1 MATSGTGVFFMLVSVCSAVSHRNQOTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
QY 61 RIHYVTFVQKMPGQLPILGEVFDSPVKGLSAFVDQPKQGAETVQGLLEKADSIIPRSHW 120
DB 61 RIHYVTFVQKMPGQLPILGEVFDSPVKGLSAFVDQPKQGAETVQGLLEKADSIIPRSHW 120
QY 121 KKTVPVLKATAGLRLPEHAKALLFEVKEIFRKSPTLVKPGSVSINDGSDGILAWTV 180

DB 121 KETPVVLTATAGLRLPEHAKALLFEVKEIFRKSPLVPKGSVIMDSDEGILAWTV 180
 QY 181 NLTQGLHGHROBTGTDLGGASTQITLPOPEKTELEOTPRGYLTSPFEMNSTYKLYTH 240
 DB 181 NLTQGLHGHROBTGTDLGGASTQITLPOPEKTELEOTPRGYLTSPFEMNSTYKLYTH 240
 QY 241 SYLGFLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGNQGEVGF 300
 DB 241 SYLGFLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGNQGEVGF 300
 QY 301 EPCYAEVLVRVGRKLHQPVEVQGSFYAFSYDDRAVDMDIDYEKGGILKVEDPERKAR 360
 DB 301 EPCYAEVLVRVGRKLHQPVEVQGSFYAFSYDDRAVDMDIDYEKGGILKVEDPERKAR 360
 QY 361 EVCNLENFTSGSPFLCMLDLSYITALLKDGFGFADSTVLQTKKVNNIETGHALGATPHL 420
 DB 361 EVCNLENFTSGSPFLCMLDLSYITALLKDGFGFADSTVLQTKKVNNIETGHALGATPHL 420
 QY 421 LQSLGISH 428
 DB 421 LQSLGISH 428

RESULT 3
 AAB72243
 ID AAB72243 standard; protein; 428 AA.
 XX AC AAB72243;
 XX DT 14-MAY-2001 (first entry)
 XX DE Human CD39 like protein CD39-L4 amino acid sequence.
 XX KW Human CD39-like protein; apyrase; NDase; platelet function inhibitor;
 KW myocardial infarction; cerebral ischaemia; angina; arterial thrombosis;
 KW cerebral artery thrombosis; platelet aggregation; inflammation;
 KW apoptosis; autoimmune disorder; neurological disorder;
 KW Alzheimer's disease; Parkinson's disease; cancer; CD39-L4.
 XX OS Homo sapiens.
 XX FN WO200110205-A1.
 XX PD 15-FEB-2001.
 XX PP 09-AUG-2000; 2000WO-US021790.
 XX PR 09-AUG-1999; 99US-00370265.
 PR 11-JAN-2000; 2000US-00481238.
 PR 25-APR-2000; 2000US-00557800.
 PR 26-MAY-2000; 2000US-00583231.
 PR 30-JUN-2000; 2000US-00608285.
 XX PA (HYSB-) HYSEQ INC.
 XX PI Ford J, Mulero JJ, Yeung G;
 XX WPI; 2001-147489/15.
 XX DR N-PSDB; AAF63402.
 XX PT Polynucleotides encoding human CD39-like polypeptides, with apyrase
 PT and/or NDase activity, which are useful in the treatment of pathological
 PT conditions caused by thrombosis (e.g. myocardial infarction) and
 PT inflammatory disorders.
 XX PS Example 9; Page 142-144; 203pp; English.
 XX CC This invention relates to polynucleotides encoding human CD39-like
 CC polypeptides with apyrase and/or NDase activity. The polypeptides having
 CC ATPase, including NDase, activity are useful for inhibiting platelet
 CC function and can therefore be used in the prophylaxis or treatment of
 CC pathological conditions caused by or involving thrombosis or excessive

CC coagulation or excessive platelet aggregation, such as myocardial
 CC infarction, cerebral ischaemia, angina, arterial thrombosis, cerebral
 CC artery thrombosis or intracardiac thrombosis, and conditions associated
 CC with venous thrombosis. CD39-L4 and CD39-L2 polypeptides are useful in
 CC modulating disease states (including platelet aggregation, inflammation
 CC and apoptosis) associated with ADP or other purinergic signalling by
 CC reducing the levels of NDPs. The polypeptides are also useful for
 CC prophylaxis or treatment of inflammation related disorders, such as
 CC disorders involving sepsis or systemic inflammatory response syndrome or
 CC SIRS (and associated conditions such as fever, tachycardia, tachypnea,
 CC cytokine overstimulation); autoimmune disorders such as thrombosis,
 CC atherosclerosis, acute pancreatitis, dermatitis, including psoriasis,
 CC cirrhosis, reperfusion injury, asthma, multiple sclerosis, arthritis;
 CC neurological disorders including neurodegenerative diseases, epilepsy,
 CC depression, Alzheimer's disease, Parkinson's disease, Huntington's
 CC disease, and amyotrophic lateral sclerosis; and cancer. The present
 CC sequence represents the CD39 like protein CD39-L4 amino acid sequence
 XX
 SQ Sequence 428 AA;
 Query Match 100.0%; Score 2250; DB 4; Length 428;
 Best Local Similarity 100.0%; Pred. No. 3.2e-208;
 Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATSWGTVFFMLVWVCVCSAVSHRNQQTWFGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
 DB 1 MATSWGTVFFMLVWVCVCSAVSHRNQQTWFGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
 QY 61 RIHVYTFVKMPGQLPILEGVFSVKPGLSAFVDQPKGARTVQGLLEVAKDSIPRSHW 120
 DB 61 RIHVYTFVKMPGQLPILEGVFSVKPGLSAFVDQPKGARTVQGLLEVAKDSIPRSHW 120
 QY 121 KKTVPVLTATAGLRLPEHAKALLFEVKEIFRKSPLVPKGSVIMDSDEGILAWTV 180
 DB 121 KKTVPVLTATAGLRLPEHAKALLFEVKEIFRKSPLVPKGSVIMDSDEGILAWTV 180
 QY 181 NLTQGLHGHROBTGTDLGGASTQITLPOPEKTELEOTPRGYLTSPFEMNSTYKLYTH 240
 DB 181 NLTQGLHGHROBTGTDLGGASTQITLPOPEKTELEOTPRGYLTSPFEMNSTYKLYTH 240
 QY 241 SYLGFLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGNQGEVGF 300
 DB 241 SYLGFLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGNQGEVGF 300
 QY 301 EPCYAEVLVRVGRKLHQPVEVQGSFYAFSYDDRAVDMDIDYEKGGILKVEDPERKAR 360
 DB 301 EPCYAEVLVRVGRKLHQPVEVQGSFYAFSYDDRAVDMDIDYEKGGILKVEDPERKAR 360
 QY 361 EVCNLENFTSGSPFLCMLDLSYITALLKDGFGFADSTVLQTKKVNNIETGHALGATPHL 420
 DB 361 EVCNLENFTSGSPFLCMLDLSYITALLKDGFGFADSTVLQTKKVNNIETGHALGATPHL 420
 QY 421 LQSLGISH 428
 DB 421 LQSLGISH 428

RESULT 4
 AAE19883
 ID AAE19883 standard; protein; 428 AA.
 XX AC AAE19883;
 XX DT 18-JUN-2002 (first entry)
 XX DE Human CD39L4 protein.
 XX KW Human CD39-like protein; CD39L4 protein; therapy; immune deficiency;
 KW autoimmune disorder; multiple sclerosis; systemic lupus erythematosus;
 KW rheumatoid arthritis; autoimmune thyroiditis; allergic reaction; asthma;
 KW insulin dependent diabetes mellitus; periodontal disease; osteoporosis;
 KW osteoarthritis; wound healing; tissue repair; Alzheimer's disease; ulcer;
 KW Parkinson's disease; amyotrophic lateral sclerosis; Huntington's disease;

nervous system disease; nerve injury; ischaemia-reperfusion injury; endotoxin lethality; arthritis; nephritis; inflammatory bowel disease; Crohn's disease; virulence; antibacterial; antiparasitic; neuroprotective; dermatological; immunosuppressive; vulnery; nootropic; anticonvulsant; antiinflammatory; nephrotropic; gastrointestinal; vasotropic.

OS Homo sapiens.
XX USG350447-B1.
XX 26-FEB-2002.
XX 29-JAN-1999; 99US-00240639.
XX 29-JAN-1999; 99US-00240639.
XX (HYSE-) HYSEQ INC.
XX Chadwick BP, Frischauf A;
XX MPI, 2002-215262/27.
XX N-PSDB; AAD31695.

An isolated polypeptide with phosphohydrolase activity, designated CD39L2, useful to identify other proteins with which binding occurs or identify inhibitors and for treatment of, e.g., Alzheimer's, multiple sclerosis and osteoporosis.

XX Example; Fig 7; 10pp; English.

XX The present invention relates to novel proteins with phosphohydrolase activity, designated CD-39-like (CD39L) proteins and polynucleotides encoding such proteins. CD39L proteins are useful to treat infectious diseases caused by viral, bacterial, fungal or other infection that may be treatable with CD39L. They are useful in the treatment of various immune deficiencies and disorders, autoimmune disorders such as multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune thyroiditis and insulin dependent diabetes mellitus, allergic reactions and conditions such as asthma and other respiratory problems, periodontal disease, osteoporosis, osteoarthritis and other tooth repair processes. They may have utility in compositions used for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration as well as for wound healing and tissue repair and replacement and in the treatment of burns, incisions and ulcers. CD39L proteins may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, i.e. for the treatment of central nervous system diseases such as Alzheimer's disease, Parkinson's disease, amyotrophic lateral sclerosis, Huntington's disease, peripheral nervous system diseases peripheral nerve injuries, peripheral neuropathy and localised neuropathies. They are also used to treat mechanical and traumatic disorders which involve degeneration, death or trauma to neural cells or nerve tissue. CD39L proteins of the invention are also useful to promote better or faster closure of non-healing wounds, including pressure ulcers, ulcers associated with vascular insufficiency and surgical and traumatic wounds. They also exhibit anti-inflammatory activity and may be used to treat inflammatory conditions including chronic or acute conditions, including ischaemia-reperfusion injury, endotoxin lethality, arthritis, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease or Crohn's disease. The present sequence is human CD39L4 protein

XX Sequence 428 AA;

Query Match 100.0%; Score 2250; DB 5; Length 428;
Best Local Similarity 100.0%; Pred. No. 3.2e-208; Indels 0; Gaps 0;
Matches 428; Conservative 0; Mismatches 0;

QY 1 MATSGTFFMLVWSCVCSAVSHRNQQTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSGTFFMLVWSCVCSAVSHRNQQTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
QY 61 RIHYTFVQKMPGQPILEGVEFDSVKGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
Db 61 RIHYTFVQKMPGQPILEGVEFDSVKGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

QY 121 KKTVPVVKATAGRLPEHAKALLPEVKEIFRKSFLVPKGSVIMDSGDEGILAWVTY 180
Db 121 KKTVPVVKATAGRLPEHAKALLPEVKEIFRKSFLVPKGSVIMDSGDEGILAWVTY 180
QY 181 NLTGQGLHGRQETVGTDLGGASTQITLPOPEKTLBOTPRGYLTSFEMNSTYKLYTH 240
Db 181 NLTGQGLHGRQETVGTDLGGASTQITLPOPEKTLBOTPRGYLTSFEMNSTYKLYTH 240
QY 241 SYLGFGKKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVYKYQYGGNOGEVGF 300
Db 241 SYLGFGKKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVYKYQYGGNOGEVGF 300
QY 301 EPCYAEVLVVRGKLHQPBEVQSGSYAFSYVDRAVDTMDIDYKGGILKVEDDFPKAR 360
Db 301 EPCYAEVLVVRGKLHQPBEVQSGSYAFSYVDRAVDTMDIDYKGGILKVEDDFPKAR 360
QY 361 EVCDNLENTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGMALGATFHL 420
Db 361 EVCDNLENTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGMALGATFHL 420
QY 421 LQSLGISH 428
Db 421 LQSLGISH 428

RESULT 5
AAY44850
ID AAY44850 standard; protein; 428 AA.
XX AC AAY44850;
XX 18-MAY-2000 (first entry)
XX Human CD39-L4 variant-ACR III mutant protein.
XX CD39-L4; human; apyrase; nucleotide diphosphatase; NDPase; variant;
XX ATP diphosphohydrolase; Arpase; adenosine diphosphate; ADP; treatment;
XX platelet aggregation; antithrombotic; thrombosis; myocardial infarction;
XX cerebral ischaemia; angina; vascular graft; extracorporeal circulation;
XX molecular weight marker; nutritional supplement; tumour; prevention;
XX drug targeting; substitution mutation.
XX Homo sapiens.
OS Synthetic.
XX Key Location/Qualifiers
XX Misc-difference 168 /note= "Wild type Asp substituted with Thr"
XX Misc-difference 170 /note= "Wild type Ser substituted with Gln"
XX Misc-difference 175 /note= "Wild type Leu substituted with Phe"
XX WO200004041-A2.
XX 27-JAN-2000.
XX 16-JUL-1999; 99WO-US016180.
XX 16-JUL-1998; 98US-00118205.
XX 24-JUL-1998; 98US-00122449.
XX 04-FEB-1999; 99US-00244444.
XX 19-MAR-1999; 99US-00273447.
XX 09-JUL-1999; 99US-00350836.
XX (HYSE-) HYSEQ INC.
XX Ford J, Mulero J;
XX WPI; 2000-182397/16.
XX N-PSDB; AAZ50357.

PT New nucleic acid encoding human CD39-like protein, useful for treating
 PT and preventing thrombotic disease.
 PS Claim 17; Fig 6; 125pp; English.
 XX
 XX The present amino acid sequence is the CD39-L4 variant, designated as ACR
 CC III mutant protein, an apyrase and/or nucleotide diphosphatase (NDPase).
 CC It is isolated from the human foetal liver-spleen cDNA library,
 CC b2HFTS20W. It is a soluble ATP diphosphohydrolase (ATPase) and is
 CC involved in the hydrolysis of adenosine diphosphate (ADP), the agonist
 CC that causes platelet aggregation. CD39-L4 protein has 30% and 80%
 CC homology to human and murine CD39. It has platelet aggregation inhibition
 CC and antithrombotic activity. CD39-L4 is used to treat or prevent
 CC thrombosis, myocardial infarction, cerebral ischaemia and angina. It is
 CC also used in vitro, to maintain vascular grafts or during extracorporeal
 CC circulation, to hydrolyse NDP, as molecular weight markers and as
 CC nutritional supplements. It is used to identify therapeutic agents that
 CC bind and modulate CD39-L4. It is coupled to toxins for targeting drugs to
 CC tumours or other cells that express CD39-L4
 XX
 XX Sequence 428 AA;
 SQ
 Query Match 99.3%; Score 2235; DB 3; Length 428;
 Best Local Similarity 99.3%; Pred. No. 9e-207;
 Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 MATSWGTVFVPMVWSCVSAVSHRNQQTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT 60
 DB 1 MATSWGTVFVPMVWSCVSAVSHRNQQTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT 60
 QY 61 RIHVYTFVQMPGQLPILEGVFDVSKFGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120
 DB 61 RIHVYTFVQMPGQLPILEGVFDVSKFGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120
 QY 121 KTFPVVLKATAGLRLLPEHKAKALLFEVKIIFRKSPLVPKGSVIMDSDEGILAWTV 180
 DB 121 KTFPVVLKATAGLRLLPEHKAKALLFEVKIIFRKSPLVPKGSVIMDSDEGILAWTV 180
 QY 121 KTFPVVLKATAGLRLLPEHKAKALLFEVKIIFRKSPLVPKGSVIMDSDEGILAWTV 180
 DB 121 KTFPVVLKATAGLRLLPEHKAKALLFEVKIIFRKSPLVPKGSVIMDSDEGILAWTV 180
 QY 181 NFLTQGLHGRQETVGTLDLGGASTQITFLPQEKLEQTPRGYLTSPFMFNSTYKLYTH 240
 DB 181 NFLTQGLHGRQETVGTLDLGGASTQITFLPQEKLEQTPRGYLTSPFMFNSTYKLYTH 240
 QY 241 SYLGFGKLAARLATLGALETGTGHTFRSACLPRMLEAWIFGGVKYQVGGNQGEVGF 300
 DB 241 SYLGFGKLAARLATLGALETGTGHTFRSACLPRMLEAWIFGGVKYQVGGNQGEVGF 300
 QY 301 EPCYAEVLVRGKIHQPEVQSGSFVAFSYVDVAVDTMDIDYKGGILKVEDPERKAR 360
 DB 301 EPCYAEVLVRGKIHQPEVQSGSFVAFSYVDVAVDTMDIDYKGGILKVEDPERKAR 360
 QY 361 EVCNDLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQITKKVANNIETGMAIGATFHL 420
 DB 361 EVCNDLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQITKKVANNIETGMAIGATFHL 420
 QY 421 LQSLGISH 428
 DB 421 LQSLGISH 428
 RESULT 6
 AAB72240
 ID AAB72240 standard; protein; 428 AA.
 XX
 XX AAB72240;
 AC
 CC
 DT 14-MAY-2001 (first entry)
 XX
 DE Human CD39 like protein CD39-L4 variant ACRIII amino acid sequence.
 XX
 KW Human CD39-like protein; apyrase; NDPase; platelet function inhibitor;
 KW myocardial infarction; cerebral ischaemia; angina; arterial thrombosis;
 KW cerebral artery thrombosis; platelet aggregation; inflammation;
 KW apoptosis; autoimmune disorder; neurological disorder; mutant; mutein;

KW Alzheimer's disease; Parkinson's disease; cancer; CD39-L4.
 XX Homo sapiens.
 XX WO200110205-A1.
 XX 15-FEB-2001.
 XX 09-AUG-2000; 2000WO-US021790.
 XX 09-AUG-1999; 99US-00370265.
 PR 11-JAN-2000; 2000US-00481238.
 PR 25-APR-2000; 2000US-00557800.
 PR 26-MAY-2000; 2000US-00583231.
 PR 30-JUN-2000; 2000US-00608285.
 XX (HYSB-) HYSEQ INC.
 PA Ford J, Mulero JJ, Yeung G;
 PI
 XX
 DR WPI; 2001-147489/15.
 DR N-PSDB; AAF63385.
 XX
 PT Polynucleotides encoding human CD39-like polypeptides, with apyrase
 PT and/or NDPase activity, which are useful in the treatment of pathological
 PT conditions caused by thrombosis (e.g. myocardial infarction) and
 PT inflammatory disorders.
 PT
 XX
 PS Claim 19; Fig 6; 203pp; English.
 XX
 CC This invention relates to polynucleotides encoding human CD39-like
 CC polypeptides with apyrase and/or NDPase activity. The polypeptides having
 CC ATPase, including NDPase, activity are useful for inhibiting platelet
 CC function and can therefore be used in the prophylaxis or treatment of
 CC pathological conditions caused by or involving thrombosis or excessive
 CC coagulation or excessive platelet aggregation, such as myocardial
 CC infarction, cerebral ischaemia, angina, arterial thrombosis, cerebral
 CC artery thrombosis or intracardiac thrombosis, and conditions associated
 CC with venous thrombosis. CD39-L4 and CD39-L2 polypeptides are useful in
 CC modulating disease states (including platelet aggregation, inflammation
 CC and apoptosis) associated with ADP or other purineergic signalling by
 CC reducing the levels of NDPs. The polypeptides are also useful for
 CC prophylaxis or treatment of inflammation related disorders, such as
 CC disorders involving sepsis or systemic inflammatory response syndrome or
 CC SIRS (and associated conditions such as fever, tachycardia, tachypnea,
 CC cytokine overstimulation); autoimmune disorders such as thrombosis,
 CC atherosclerosis, acute pancreatitis, dermatitis, including psoriasis,
 CC cirrhosis, reperfusion injury, asthma, multiple sclerosis, arthritis;
 CC neurological disorders including neurodegenerative diseases, epilepsy,
 CC depression, Alzheimer's disease, Parkinson's disease, Huntington's
 CC disease, and amyotrophic lateral sclerosis; and cancer. The present
 CC sequence represents human CD39 like protein CD39-L4 variant ACRIII
 XX
 SQ Sequence 428 AA;
 Query Match 99.3%; Score 2235; DB 4; Length 428;
 Best Local Similarity 99.3%; Pred. No. 9e-207;
 Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 MATSWGTVFVPMVWSCVSAVSHRNQQTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT 60
 DB 1 MATSWGTVFVPMVWSCVSAVSHRNQQTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT 60
 QY 61 RIHVYTFVQMPGQLPILEGVFDVSKFGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120
 DB 61 RIHVYTFVQMPGQLPILEGVFDVSKFGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120
 QY 121 KTFPVVLKATAGLRLLPEHKAKALLFEVKIIFRKSPLVPKGSVIMDSDEGILAWTV 180
 DB 121 KTFPVVLKATAGLRLLPEHKAKALLFEVKIIFRKSPLVPKGSVIMDSDEGILAWTV 180
 QY 181 NFLTQGLHGRQETVGTLDLGGASTQITFLPQEKLEQTPRGYLTSPFMFNSTYKLYTH 240
 DB 181 NFLTQGLHGRQETVGTLDLGGASTQITFLPQEKLEQTPRGYLTSPFMFNSTYKLYTH 240

Db 181 NLTGQLHGRQETVGLDGGASTQITLPPQFKTLTQTPRGYLTSEFNFSTYKLYTH 240
Qy 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGVKYQYGGNQGEVGF 300
Db 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGVKYQYGGNQGEVGF 300
Qy 301 EPCVAEVLVRVGRKHOPEEVORGSFVAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 301 EPCVAEVLVRVGRKHOPEEVORGSFVAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
Qy 361 EVCNLENFTSGSPFLCMLDLSYITALLKDGFGFADSTVLQLTKVNNIETGWLGAFFHL 420
Db 361 EVCNLENFTSGSPFLCMLDLSYITALLKDGFGFADSTVLQLTKVNNIETGWLGAFFHL 420
Qy 421 LOSLGISH 428
Db 421 LOSLGISH 428

RESULT 7

AAV44851

ID AAV44851 standard; protein; 405 AA.

XX AC AAV44851;

XX DT 18-MAY-2000 (first entry)

XX DE Human CD39-L66 protein.

XX CD39-L66; human; CD39-L66; apyrase; nucleotide diphosphatase; NDPase;
KW ATP diphosphohydrolase; ATPase; adenosine diphosphate; ADP; treatment;
KW platelet aggregation; antithrombotic; thrombosis; myocardial infarction;
KW cerebral ischaemia; angina; vascular graft; extracorporeal circulation;
KW molecular weight marker; nutritional supplement; tumour; prevention;
KW drug targeting; splice variant.

XX OS Homo sapiens.

XX PN WO200004041-A2.

XX PD 27-JAN-2000.

XX PF 16-JUL-1999; 99WO-US016180.

XX PR 16-JUL-1999; 98US-00118205.

XX PR 24-JUL-1998; 98US-00122449.

XX PR 04-FEB-1999; 99US-0024444.

XX PR 19-MAR-1999; 99US-00273447.

XX PR 09-JUL-1999; 99US-00350836.

XX PA (HYSE-) HYSEQ INC.

XX PI Ford J, Mulero J;

XX DR WPI; 2000-182397/16.

XX DR N-PSDB; AAZ50358.

XX PT New nucleic acid encoding human CD39-like protein, useful for treating
XX and preventing thrombotic disease.
XX Claim 15; Page 124-125; 125pp; English.

XX The present amino acid sequence is the CD39-L66 protein, a splice variant
XX of the CD39-L4 protein. It is an apyrase and/or nucleotide diphosphatase
XX (NDPase), isolated from the human foetal liver-spleen cDNA library,
XX b2HLIS20W. It is a soluble ATP diphosphohydrolase (ATPDase) and is
XX involved in the hydrolysis of adenosine diphosphate (ADP), the agonist
XX that causes platelet aggregation. CD39-L4 protein has 30% and 80%
XX homology to human and murine CD39. It has platelet aggregation inhibition
XX and antithrombotic activity. CD39-L4 is used to treat or prevent
XX thrombosis, myocardial infarction, cerebral ischaemia and angina. It is
XX also used in vitro, to maintain vascular grafts or during extracorporeal
XX circulation, to hydrolyse NDP, as molecular weight markers and as

CC nutritional supplements. It is used to identify therapeutic agents that
CC bind and modulate CD39-L4. It is coupled to toxins for targeting drugs to
CC tumours or other cells that express CD39-L4

XX SQ Sequence 405 AA;

Query Match 93.5%; Score 2104; DB 3; Length 405;

Best Local Similarity 100.0%; Pred. No. 3.7e-194;

Matches 400; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATSNCTVFFMLVWSCVSAVSHRNQQTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Db 1 MATSNCTVFFMLVWSCVSAVSHRNQQTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy 61 RIHVYTFVQKMPQOLPILEGEVFDVSKPGLSAFVDPQKGAETVQGLLEVAKDSIPRSHW 120

Db 61 RIHVYTFVQKMPQOLPILEGEVFDVSKPGLSAFVDPQKGAETVQGLLEVAKDSIPRSHW 120

Qy 121 KKTPTVVLKATAGLRLPEHAKALLPEVKEIIFRKSPLFVPGKSVSMDSDGILAWTV 180

Db 121 KKTPTVVLKATAGLRLPEHAKALLPEVKEIIFRKSPLFVPGKSVSMDSDGILAWTV 180

Qy 181 NFLTGLHGRQETVGLDGGASTQITLPPQFKTLTQTPRGYLTSEFNFSTYKLYTH 240

Db 181 NFLTGLHGRQETVGLDGGASTQITLPPQFKTLTQTPRGYLTSEFNFSTYKLYTH 240

Qy 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGVKYQYGGNQGEVGF 300

Db 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGVKYQYGGNQGEVGF 300

Qy 301 EPCVAEVLVRVGRKHOPEEVORGSFVAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360

Db 301 EPCVAEVLVRVGRKHOPEEVORGSFVAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360

Qy 361 EVCNLENFTSGSPFLCMLDLSYITALLKDGFGFADSTVLQ 400

Db 361 EVCNLENFTSGSPFLCMLDLSYITALLKDGFGFADSTVLQ 400

RESULT 8

AAV72239

ID AAV72239 standard; protein; 405 AA.

XX AC AAV72239;

XX DT 14-MAY-2001 (first entry)

XX DE Human CD39 like protein CD39-L4 amino acid sequence.

XX KW Human CD39-like protein; apyrase; NDPase; platelet function inhibitor;
XX myocardial infarction; cerebral ischaemia; angina; arterial thrombosis;
XX cerebral artery thrombosis; platelet aggregation; inflammation;
XX apoptosis; autoimmune disorder; neurological disorder;
XX Alzheimer's disease; Parkinson's disease; cancer; CD39-L4.

XX OS Homo sapiens.

XX PN WO200110205-A1.

XX PD 15-FEB-2001.

XX PF 09-AUG-2000; 2000WO-US021790.

XX PR 09-AUG-1999; 99US-00370265.

XX PR 11-JAN-2000; 2000US-00481238.

XX PR 25-APR-2000; 2000US-00557800.

XX PR 26-MAY-2000; 2000US-00583231.

XX PR 30-JUN-2000; 2000US-00608285.

XX PA (HYSE-) HYSEQ INC.

XX PI Ford J, Mulero JJ, Yeung G;

XX

DR WPI: 2001-147489/15.
PT N-PSDB; AAF63384.
XX Polynucleotides encoding human CD39-like polypeptides, with apyrase
PT and/or NTPase activity, which are useful in the treatment of pathological
PT conditions caused by thrombosis (e.g. myocardial infarction) and
PT inflammatory disorders.
XX
PS
XX Claim 17: Page 157-158; 203pp; English.
XX
XX This invention relates to polynucleotides encoding human CD39-like
CC polypeptides with apyrase and/or NTPase activity. The polypeptides having
CC APTase, including NTPase, activity are useful for inhibiting platelet
CC function and can therefore be used in the prophylaxis or treatment of
CC pathological conditions caused by or involving thrombosis or excessive
CC coagulation or excessive platelet aggregation, such as myocardial
CC infarction, cerebral ischemia, angina, arterial thrombosis, cerebral
CC artery thrombosis or intracardiac thrombosis and conditions associated
CC with venous thrombosis. CD39-L4 and CD39-L2 polypeptides are useful in
CC modulating disease states (including platelet aggregation, inflammation
CC and apoptosis) associated with ADP or other purinergic signalling by
CC reducing the levels of NTPs. The polypeptides are also useful for
CC prophylaxis or treatment of inflammation related disorders, such as
CC disorders involving sepsis or systemic inflammatory response syndrome or
CC SIRS (and associated conditions such as fever, tachycardia, tachypnea,
CC cytokine overstimulation); autoimmune disorders such as thrombosis,
CC atherosclerosis, acute pancreatitis, dermatitis, including psoriasis,
CC cirrhosis, reperfusion injury, asthma, multiple sclerosis, arthritis;
CC neurological disorders including neurodegenerative diseases, epilepsy,
CC depression, Alzheimer's disease, Parkinson's disease, Huntington's
CC disease, and amyotrophic lateral sclerosis; and cancer. The present
CC sequence represents human CD39 like protein CD39-L4
XX
XX Sequence 405 AA;

Query Match 93.5%; Score 2104; DB 4; Length 405;
Best Local Similarity 100.0%; Pred. No. 3.7e-194;
Matches 400; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MATSWGTVPMVYVSCVSNVSHRNOOTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Db |||||
Qy 61 RIHYTVFQVMPQLPLEGEVDSVKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120
Db |||||
Qy 61 RIHYTVFQVMPQLPLEGEVDSVKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120
Db |||||
Qy 121 KKTPTVVKATAGRLLEPHKAKALLFEVKEIPKSPFLPKGSVIMDSGDEGILAWTV 180
Db |||||
Qy 121 KKTPTVVKATAGRLLEPHKAKALLFEVKEIPKSPFLPKGSVIMDSGDEGILAWTV 180
Db |||||
Qy 181 NPLTGLHGRHQRQETVGTDLGGASTQITPLPQEKLETPPGYLTSPFENSTYKLYTH 240
Db |||||
Qy 181 NPLTGLHGRHQRQETVGTDLGGASTQITPLPQEKLETPPGYLTSPFENSTYKLYTH 240
Db |||||
Qy 241 SYLGFGKAAARLATLGALETGDTGHTFRSACLPRWLEAWIPFGVKYQYGGNQEVEGF 300
Db |||||
Qy 241 SYLGFGKAAARLATLGALETGDTGHTFRSACLPRWLEAWIPFGVKYQYGGNQEVEGF 300
Db |||||
Qy 301 EPCYAEVLVRGKLHPREVORGSFVASYYYDRAVDMDIDYKGGILKVEDFERKAR 360
Db |||||
Qy 301 EPCYAEVLVRGKLHPREVORGSFVASYYYDRAVDMDIDYKGGILKVEDFERKAR 360
Db |||||
Qy 361 EVCNLENFTSGSPFLCLMDLSYITALLKDGFGFADSTVLQ 400
Db |||||
Qy 361 EVCNLENFTSGSPFLCLMDLSYITALLKDGFGFADSTVLQ 400
Db |||||

RESULT 9
AAE19884
ID AAE19884 standard; protein; 465 AA.
XX
AC AAE19884;

XX DT 18-JUN-2002 (first entry)
XX DE Mouse CD39L4 protein.
XX KW Mouse CD39L4 protein.
KW Mouse; CD-39-like protein; CD39L4 protein; therapy; immune deficiency;
KW autoimmune disorder; multiple sclerosis; systemic lupus erythematosus;
KW rheumatoid arthritis; autoimmune thyroiditis; allergic reaction; asthma;
KW insulin dependent diabetes mellitus; periodontal disease; osteoporosis;
KW osteoarthritis; wound healing; tissue repair; Alzheimer's disease; ulcer;
KW Parkinson's disease; amyotrophic lateral sclerosis; Huntington's disease;
KW nervous system disease; nerve injury; ischaemia-reperfusion injury;
KW endotoxin lethality; arthritis; nephritis; inflammatory bowel disease;
KW Crohn's disease; viricide; antibacterial; antifungal; neuroprotective;
KW dermatological; immunosuppressive; vulnery; nootropic; anticonvulsant;
KW antiinflammatory; nephrotropic; gastrointestinal; vasotropic; NTPase;
KW nucleotide-triphosphatase; enzyme.
XX
OS Mus musculus.
XX US6350447-B1.
PN 26-FEB-2002.
XX 29-JAN-1999; 99US-00240639.
XX 29-JAN-1999; 99US-00240639.
PD (HYSE-) HYSEQ INC.
XX Chadwick BP, Frieschaut A;
XX WPI: 2002-215262/27.
DR N-PSDB; AAD31696.
XX
XX An isolated polypeptide with phosphohydrolase activity, designated
PT CD39L2, useful to identify other proteins with which binding occurs or
PT identify inhibitors and for treatment of, e.g., Alzheimer's, multiple
PT sclerosis and osteoporosis.
XX
PS Example; Fig 1; 101pp; English.

The present invention relates to novel proteins with phosphohydrolase activity, designated CD-39-like (CD39L) proteins and polynucleotides encoding such proteins. CD39L proteins are useful to treat infectious diseases caused by viral, bacterial, fungal or other infection that may be treatable with CD39L. They are useful in the treatment of various immune deficiencies and disorders, autoimmune disorders such as multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune thyroiditis and insulin dependent diabetes mellitus, allergic reactions and conditions such as asthma and other respiratory problems, periodontal disease, osteoporosis, osteoarthritis and other tooth repair processes. They may have utility in compositions used for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration as well as for wound healing and tissue repair and replacement and in the treatment of burns, incisions and ulcers. CD39L proteins may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, i.e. for the treatment of central nervous system diseases such as Alzheimer's disease, Parkinson's disease, amyotrophic lateral sclerosis, Huntington's disease, peripheral nervous system diseases peripheral nerve injuries, peripheral neuropathy and localised neuropathies. They are also used to treat mechanical and traumatic disorders which involve degeneration, death or trauma to neural cells or nerve tissue. CD39L proteins of the invention are also useful to promote better or faster closure of non-healing wounds, including pressure ulcers, ulcers associated with vascular insufficiency and surgical and traumatic wounds. They also exhibit anti-inflammatory activity and may be used to treat inflammatory conditions including chronic or acute conditions), including ischaemia-reperfusion injury, endotoxin lethality, arthritis, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease or Crohn's disease. The present sequence is mouse CD39L4 protein, also known as nucleotide-triphosphatase (NTPase)

XX PS Claim 11; Page 1429-1431; 2104pp; English.

XX CC AAC97991 to AAC98763 encode the human colon cancer associated proteins, called human colon cancer antigens, given in AAB53234 to AAB54006. The human colon cancer antigens can have cytostatic, cardioactive, muscular; neuroprotective, immunomodulatory, gynaecological, gastrointestinal, and vulnary, nephrotropic, antiinfective and antibacterial activities, and can be used in gene therapy. The colon cancer antigen polynucleotides, proteins and antibodies to the proteins are useful for the prevention, treatment and diagnosis of colon disorders, such as colon cancer. The polynucleotides may be used in diagnostics and research, such as for chromosome identification, and as hybridisation probes. The proteins may also be used to prevent diseases such as neural disorders, immune system disorders, muscular disorders, reproductive disorders, gastrointestinal disorders, wounds, renal disorders, infectious diseases, and cardiovascular disorders. AAC98764 to AAC98772 and AAB54007 represent sequences used in the exemplification of the present invention

XX SQ Sequence 330 AA;

Query Match 73.8%; Score 1660; DB 3; Length 330;
Best Local Similarity 97.3%; Pred. No. 2.1e-151;
Matches 319; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 101 AETVQGLLEVAKDSIPRSHWKTTPVVLKATAGLRLPPEHKKALLFEVKGFPRKSPFLVP 160
DB 3 ARVQGLLEVAKDSIPRSHWKTTPVVLKATAGLRLPPEHKKALLFEVKGFPRKSPFLVP 62

QY 161 KGSVIMDGSDEGILAWTVNLTGQLHGHROSTVCTLDLGGASTQTITLPOPEKTLQET 220
DB 63 KGSVIMDGSDEGILAWTVNLTGQLHGHROSTVCTLDLGGASTQTITLPOPEKTLQET 122

QY 221 PRGVLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETGDTGDTFRSACLPRWLEAE 280
DB 123 PXGVLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETGDTGDTFRSACLPRWLEAE 182

QY 281 WIFGVKYQYQGNQGEVGPVCAEVLVVRGKLHQPVEVQSGSFYAFSVYDRAVDTD 340
DB 183 WIFGVKYQYQGNQGEVGPVCAEVLVVRGKLHQPVEVQSGSFYAFSVYDRAVDTD 242

QY 341 MIDYKGGILKVEDFERKAREVCDNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQ 400
DB 243 MIDYKGGILKVEDFERKAREVCDNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQ 302

QY 401 LTKVNNIETGVALGATFHLLQSLGISH 428
DB 303 LTKVNNIETGVALGATFHLLQSLGISH 330

RESULT 11
ID AAE19881 standard; protein; 456 AA.
XX AAE19881;
XX AC AAE19881;
XX DT 18-JUN-2002 (first entry)
XX DE Human CD39L2 protein.
XX KW Human; CD-39-like protein; CD39L2 protein; therapy; immune deficiency; autoimmune disorder; multiple sclerosis; systemic lupus erythematosus; rheumatoid arthritis; autoimmune thyroiditis; allergic reaction; asthma; insulin dependent diabetes mellitus; periodontal disease; osteoporosis; osteoarthritis; wound healing; tissue repair; Alzheimer's disease; ulcer; Parkinson's disease; amyotrophic lateral sclerosis; Huntington's disease; nervous system disease; nerve injury; ischaemia-reperfusion injury; endotoxin lethality; arthritis; nephritis; inflammatory bowel disease; Crohn's disease; virucide; antibacterial; antifungal; neuroprotective; dermatological; immunosuppressive; vulnary; nephrotropic; anticonvulsant; antiinflammatory; nephrotropic; gastrointestinal; vasotropic.

XX OS Homo sapiens.

SQ Sequence 465 AA;

Query Match 81.4%; Score 1832.5; DB 5; Length 465;
Best Local Similarity 83.8%; Pred. No. 7.9e-168;
Matches 352; Conservative 25; Mismatches 40; Indels 3; Gaps 3;

QY 1 MATSGTVPFVNLVSCVSAVSHRNQQTWPGIFLSSMCPINVSASTLYGIMFAGSTGT 60
DB 1 MATSGAV-FMLIIACVGTVPFVNLVSCVSAVSHRNQQTWPGIFLSSMCPINVSAGTFYIMFAGSTGA 59

QY 61 RIHVYFVQKMPGQLPILEGVFDVSKPGLSAFVDQPKQAGTVOGLLEVAKDSIPRSHW 120
DB 60 RIHVYFVQKMPGQLPILEGVFDVSKPGLSAFVDQPKQAGTVOGLLEVAKDSIPRSHW 119

QY 121 KKTVPVLKATAGLRLPPEHKKALLFEVKGFPRKSPFLVPKGSVIMDGSDEGILAWTV 180
DB 120 KKTVPVLKATAGLRLPPEHKKALLFEVKGFPRKSPFLVPKGSVIMDGSDEGILAWTV 179

QY 181 NPLTGQLHGHROSTVCTLDLGGASTQTITLPOPEKTLQETPRGVLTSFEMFNSTYKLYTH 240
DB 180 NPLTGQLHGHROSTVCTLDLGGASTQTITLPOPEKTLQETPRGVLTSFEMFNSTYKLYTH 239

QY 241 SYLGFGLKAARLATLGALETGDTGDTFRSACLPRWLEAEWIFGVKYQYQGNQGEVGP 300
DB 240 SYLGFGLKAARLATLGALETGDTGDTFRSACLPRWLEAEWIFGVKYQYQGNQGEVGP 299

QY 301 BPCYAEVLVVRGKLHQPVEVQSGSFYAFSVYDRAVDTDMDYKGGILKVEDFERKAR 360
DB 300 BPCYAEVLVVRGKLHQPVEVQSGSFYAFSVYDRAVDTDMDYKGGILKVEDFERKAR 359

QY 361 EYVCDNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTLTKVNNIETGVALGATFHL 420
DB 360 EYVCDNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTLTKVNNIETGVALGATFHL 417

RESULT 10
ID AAB53336 standard; protein; 330 AA.
XX AAB53336;
XX AC AAB53336;
XX DT 09-MAR-2001 (first entry)
XX DE Human colon cancer antigen protein sequence SEQ ID NO:876.
XX KW Human; colon cancer; colon cancer antigen; diagnosis; detection; identification; cytostatic; cardioactive; neuroprotective; vulnary; immunomodulatory; muscular; gynaecological; gastrointestinal; nephrotropic; antiinfective; antibacterial; gene therapy; wound; neural disorder; immune system disorder; muscular disorder; reproductive disorder; gastrointestinal disorder; renal disorder; infectious disease; cardiovascular disorder.
XX OS Homo sapiens.
XX PN WO200055351-A1.
XX PD 21-SEP-2000.
XX PF 08-MAR-2000; 2000WO-US005883.
XX PR 12-MAR-1999; 99US-0124270P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Ruben SM;
XX DR WPI; 2000-587534/55.
XX DR N-PSDB; AAC98093.
XX KW Colon cancer associated gene sequences, referred to as colon cancer antigens, useful for the treatment, prevention, and diagnosis of colon disorders such as colon cancer.

[illegible]

241	QY	DB	LOASEPPGVLTALRNFNRNTYKLYSYSLGCLGMARLAILGGVEGQAPCKOCKELVSPCLSP	300
276	QY	DB	WLEAEWIFGGVKYQYGGNQBCEVGFEPFCYAEVLVRVRGKLHQPEVRQSGFYAFSYYDYR	335
301	DB	DB	SPFGWEHAETVYRVSGQKAAASLHELCAARVSEVLQNRVHRTTEVKHVDIFYAFSYYDL	360
336	QY	DB	AVTDMIDYERGGILKVEDFERKAREVCDNLENFTSGSPFLCMLDSYITALLKDGCGPAD	395
361	DB	DB	AAGVGLIDAEKGGSLVGDPEFAIAAKYVCRITLTPQSSPSPSCMDLTVSVLLQF-FGPPR	419
396	QY	DB	STVLOITKKKNITGHALGATFFHLQSL	424
420	DB	DB	SKVLKLRKIDNVETSWALGAIFYIDSL	448

RESULT 12	
AAB72241	standard; protein; 484 AA.
ID	
XX	
AC	AAB72241;
XX	
DT	14-MAY-2001 (first entry)
XX	
DE	Human CD39 like protein CD39-L2 amino acid sequence.
XX	
KW	Human CD39-like protein; apyrase; NDPAse; platelet function inhibitor;
KW	myocardial infarction; cerebral ischaemia; angina; arterial thrombosis;
KW	cerebral artery thrombosis; platelet aggregation; inflammation;
KW	apoptosis; autoimmune disorder; neurological disorder;
KW	Alzheimer's disease; Parkinson's disease; cancer; CD39-L2.
XX	
OS	Homo sapiens.
XX	
PN	WO200110205-A1.
XX	
PD	15-FEB-2001.
XX	
PF	09-AUG-2000; 2000WO-US021790.
XX	
PR	09-AUG-1999; 99US-00370265.
PR	11-JAN-2000; 2000US-00481238.
PR	25-APR-2000; 2000US-0057800.
PR	26-MAY-2000; 2000US-00583231.
PR	30-JUN-2000; 2000US-00608285.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
FI	Ford J, Mulero JJ, Yeung G;
XX	
DR	WPI; 2001-147489/15.
DR	N-PSDB: AAP63386.

polynucleotides encoding human CD39-like polypeptides, with apyrase and/or NDPase activity, which are useful in the treatment of pathological conditions caused by thrombosis (e.g. myocardial infarction) and inflammatory disorders.

Claim 39; Page 162-164; 203pp; English.

This invention relates to polynucleotides encoding human CD39-like polypeptides with apyrase and/or NDPase activity. The polypeptides having ATPdase, including NDPase, activity are useful for inhibiting platelet function and can therefore be used in the prophylaxis or treatment of pathological conditions caused by or involving thrombosis or excessive coagulation or excessive platelet aggregation, such as myocardial infarction, cerebral ischaemia, angina, arterial thrombosis, cerebral artery thrombosis or intracardiac thrombosis, and conditions associated with venous thrombosis. CD39-14 and CD39-12 polypeptides are useful in modulating disease states (including platelet aggregation, inflammation and apoptosis) associated with ADP or other purinergic signalling by reducing the levels of NDPs. The polypeptides are also useful for prophylaxis or treatment of inflammation related disorders, such as disorders involving sepsis or systemic inflammatory response syndrome or

	Query Match	44.4%	Score 999;	DB 5;	Length 456;
	Best Local Similarity	52.4%;	Pred. No. 2.6e-87;		
	Matches 204; Conservative	56;	Mismatches 123;	Indels	6; Gaps
Qy	40 PINVSA--STLYGIMPDAGSGTGRHHVTYVQRKMPGLPILEGVEDSDVKPGLSFAVDQ	96			
	: : : : : : : : : : : : : : : :				
Dd	62 PLGTAAADGHEVFYGYGIMFPDAGSGTGTRHVHFQT-RPPRETPLTHETFKAVKEGLSAYADD	120			
	: : : : : : : : : : : : : : : :				
Qy	97 PKQGAETVOGLELVAKDSISPRSHWKKTPVLVKATAGLRLLPEHKAKALLPEVKSIIPRKSP	156			
	: : : : : : : : : : : : : : : :				
Dd	121 VESAQGIRELLDVAKODI-PDFFWKA-PTVLVKATAGLRLLPEHKAQKLLOKVKEVFRASP	180			
	: : : : : : : : : : : : : : : :				
Qy	157 FLVPKGVSVMNDSDEGILAWVNFLTQLGHRRQETGVLTDLGGASTQTITFLPOEFT	216			
	: : : : : : : : : : : : : : : :				
Dd	181 FLVGDDCVSIMNGTDEGSAMITINFLTGSTLKTFGGWSGMLDGGGSTQTAF-LRVREGT	240			
	: : : : : : : : : : : : : : : :				
Qy	217 LEOTPGRLTSFEMENSTKYLTYSYLGFGLKAARLATIGLALETB-GTDGHTFRSACLPR	275			

CC chronic diseases, and cancer. The isolated protein, its encoding
CC polynucleotide or an antibody created from the protein are useful for
CC treating or preventing neurological disorders such as epilepsy, stroke,
CC mental disorders including mood, anxiety, schizophrenic disorders,
CC disorders of vesicular transport such as cystic fibrosis, diabetes
CC mellitus, goiter, gastrointestinal disorders including ulcerative
CC colitis, other conditions associated with abnormal vesicle trafficking
CC including AIDS, allergic reactions, multiple sclerosis and rheumatoid
CC arthritis. A cell comprising the vector of the invention is useful for
CC producing non-human transgenic animals. The polynucleotide of the
CC invention can be used to treat disorders by gene therapy. This sequence
CC represents one of the isolated NOVX proteins of the invention
XX
XX Sequence 467 AA;
SQ
Query Match 44.3%; Score 996; DB 5; Length 467;
Best Local Similarity 52.2%; Pred. No. 5.2e-87;
Matches 203; Conservative 57; Mismatches 123; Indels 6; Gaps 4;
QY 40 PINVA---STLYGIMFDAGSTGTHVTVFVQKMPGQLPILGEVEDSVKPLSAFVDQ 96
Db 73 PLGTAADGHEVYFIMFDAGSTGTHVTVFVQKMPGQLPILGEVEDSVKPLSAFVDQ 131
QY 97 PRQGAETVQGLLEVARQDSIPRSHWKTTPVLKATAGLRLLEPHKAKALLFEVKEIFRKSP 156
Db 132 VEKSAGGIRELLDVAQDIPDFWKATPLVLKATAGLRLLEPHKAKALLFEVKEIFRKSP 191
QY 157 FLVPKGSVIMDSDEGILLAVTVNPLTQGLHGRQETVGLDLGGCASTOITLPOPEKT 216
Db 192 FLVGDDCVSINMGTDGVSAMITINFLTGLSKTPGGSSVGMDDGGSGTQIAELPRVEGT 251
QY 217 LEOTPRGYLTSPFEMFNSTVLYTHSYLGFGLKAARLATIAGLETE-GTIDHTFRSACLPR 275
Db 252 LQASPPGYLTALFMFRTYKLYSYLGLGLNARLAILGGVGPQAKDKGLVSPCLSP 311
QY 276 WLEAEWIFGQVYQYCGNORGEVGFPCYAEVLVRVGRKLGHPQEEVQSGFYAFSYVDYR 335
Db 312 SPKGEWEHAEVTVRVSGQKAAASHELCAARVSEVLQNRVHRTVEEVKGVDFYAFSYVDL 371
QY 336 AVTDMIDYKGGIILKVEDFERKAREVCNLENFTSGSPFLCMDLSYITALLKDGFGFAD 395
Db 372 AAGVGUIDAEKGGSLVVGDFEIAAKVVCRTLETPQSSPFCMDLTVVSLLLQE-FGFP 430
QY 396 STYQLTKKYNNTETGALGATFHLLOSL 424
Db 431 SKVLKTRKIDNVETSWALGAIFFHYIDSL 459
RESULT 15
ID AAM93929
XX AAM93929 standard; protein; 456 AA.
XX AC AAM93929;
XX DT 06-NOV-2001 (first entry)
XX DE Human polypeptide, SEQ ID NO: 4100.
XX KW Human; full length cDNA; cDNA synthesis; oligo-capping.
XX OS Homo sapiens.
XX PN EP1130094-A2.
XX PD 05-SEP-2001.
XX PF 07-JUL-2000; 2000EP-00114089.
XX PR 08-JUL-1999; 99JP-00194486.
XX PR 11-JAN-2000; 2000JP-00118774.
XX PR 02-MAY-2000; 2000JP-00183765.
XX PA (HELI-) HELIX RES INST.

CC chronic diseases, and cancer. The isolated protein, its encoding
CC polynucleotide or an antibody created from the protein are useful for
CC treating or preventing neurological disorders such as epilepsy, stroke,
CC mental disorders including mood, anxiety, schizophrenic disorders,
CC disorders of vesicular transport such as cystic fibrosis, diabetes
CC mellitus, goiter, gastrointestinal disorders including ulcerative
CC colitis, other conditions associated with abnormal vesicle trafficking
CC including AIDS, allergic reactions, multiple sclerosis and rheumatoid
CC arthritis. A cell comprising the vector of the invention is useful for
CC producing non-human transgenic animals. The polynucleotide of the
CC invention can be used to treat disorders by gene therapy. This sequence
CC represents one of the isolated NOVX proteins of the invention
XX
XX Sequence 467 AA;
SQ
Query Match 44.3%; Score 996; DB 5; Length 467;
Best Local Similarity 52.2%; Pred. No. 5.2e-87;
Matches 203; Conservative 57; Mismatches 123; Indels 6; Gaps 4;
QY 40 PINVA---STLYGIMFDAGSTGTHVTVFVQKMPGQLPILGEVEDSVKPLSAFVDQ 96
Db 73 PLGTAADGHEVYFIMFDAGSTGTHVTVFVQKMPGQLPILGEVEDSVKPLSAFVDQ 131
QY 97 PRQGAETVQGLLEVARQDSIPRSHWKTTPVLKATAGLRLLEPHKAKALLFEVKEIFRKSP 156
Db 132 VEKSAGGIRELLDVAQDIPDFWKATPLVLKATAGLRLLEPHKAKALLFEVKEIFRKSP 191
QY 157 FLVPKGSVIMDSDEGILLAVTVNPLTQGLHGRQETVGLDLGGCASTOITLPOPEKT 216
Db 192 FLVGDDCVSINMGTDGVSAMITINFLTGLSKTPGGSSVGMDDGGSGTQIAELPRVEGT 251
QY 217 LEOTPRGYLTSPFEMFNSTVLYTHSYLGFGLKAARLATIAGLETE-GTIDHTFRSACLPR 275
Db 252 LQASPPGYLTALFMFRTYKLYSYLGLGLNARLAILGGVGPQAKDKGLVSPCLSP 311
QY 276 WLEAEWIFGQVYQYCGNORGEVGFPCYAEVLVRVGRKLGHPQEEVQSGFYAFSYVDYR 335
Db 312 SPKGEWEHAEVTVRVSGQKAAASHELCAARVSEVLQNRVHRTVEEVKGVDFYAFSYVDL 371
QY 336 AVTDMIDYKGGIILKVEDFERKAREVCNLENFTSGSPFLCMDLSYITALLKDGFGFAD 395
Db 372 AAGVGUIDAEKGGSLVVGDFEIAAKVVCRTLETPQSSPFCMDLTVVSLLLQE-FGFP 430
QY 396 STYQLTKKYNNTETGALGATFHLLOSL 424
Db 431 SKVLKTRKIDNVETSWALGAIFFHYIDSL 459
RESULT 15
ID AAM93929
XX AAM93929 standard; protein; 456 AA.
XX AC AAM93929;
XX DT 06-NOV-2001 (first entry)
XX DE Human polypeptide, SEQ ID NO: 4100.
XX KW Human; full length cDNA; cDNA synthesis; oligo-capping.
XX OS Homo sapiens.
XX PN EP1130094-A2.
XX PD 05-SEP-2001.
XX PF 07-JUL-2000; 2000EP-00114089.
XX PR 08-JUL-1999; 99JP-00194486.
XX PR 11-JAN-2000; 2000JP-00118774.
XX PR 02-MAY-2000; 2000JP-00183765.
XX PA (HELI-) HELIX RES INST.

DB 427 SKVLKTRKIDNVETSWALGAIFFHYIDSL 455
RESULT 14
ABJ04657
ID ABJ04657 standard; protein; 467 AA.
XX AC ABJ04657;
XX DT 11-OCT-2002 (first entry)
XX DE Protein of NOVX 15a SEQ ID No 36.
XX KW Cytostatic; antidiabetic; anorectic; metabolic; nootropic; antilipaemic;
KW neuroprotective; antiparkinsonian; anticonvulsant; cerebroprotective;
KW tranquiliser; neuroleptic; antidiabetic; antitumor; antiinflammatory;
KW anti-HIV; antiallergic; antirheumatic; antiarthritic; NOVX; diabetes;
KW metabolic disorder; obesity; infectious diseases; Alzheimer's disease;
KW anorexia; neurodegenerative disorder; Parkinson's disorder; obesity;
KW immune disorder; haematopoietic disorder; dyslipidaemia; chronic disease;
KW metabolic syndrome X; wasting disorder; cancer; neurological disorder;
KW epilepsy; stroke; mental disorder; schizophrenic disorders; goiter;
KW vesicular transport; cystic fibrosis; gastrointestinal disorder;
KW diabetes mellitus; ulcerative colitis; AIDS; allergic reaction;
KW multiple sclerosis; rheumatoid arthritis; transgenic animal;
KW gene therapy.
XX OS Unidentified.
XX PN W0200246409-A2.
XX PD 13-JUN-2002.
XX PF 06-DEC-2001; 2001WO-US046586.
XX PR 06-DEC-2000; 2000US-0251660P.
XX PR 12-DEC-2000; 2000US-0255029P.
XX PR 08-JAN-2001; 2001US-0260326P.
XX PR 24-JAN-2001; 2001US-0263800P.
XX PR 20-FEB-2001; 2001US-0269942P.
XX PR 24-APR-2001; 2001US-0286183P.
XX PR 20-AUG-2001; 2001US-0313627P.
XX PR 12-SEP-2001; 2001US-0318712P.
XX PA (CURA-) CURAGEN CORP.
XX PI Guo X, Li L, Patturajan M, Shamkats RA, Casman SJ, Malyankar UM;
PI Tchernev VT, Vernet CM, Spytek KA, Shenoy SG, Alsobrook JP;
PI Edinger S, Peyman JA, Stone DJ, Ellerman K, Gangolli EA, Boldog FL;
PI Colman SD, Eisen AJ, Liu X, Padigar M, Spaderna SK, Zerhusen BD;
XX WPI; 2002-547774/58.
XX DR N-PSDB; AET05470.
XX PT Novel isolated polypeptide, designated NOVX, useful for treating or
XX preventing cancer, diabetes, obesity, dyslipidemia, anorexia, and
XX metabolic, neurodegenerative, immune and hematopoietic disorders.
XX FS Claim 1; Page 140; 421pp; English.
XX CC The invention relates to an isolated polypeptide, designated NOVX,
XX comprising a sequence fully defined in the specification. The isolated
XX protein, its encoding polynucleotide or an antibody created from the
XX protein is useful in the manufacture of a medicament for treating a
XX syndrome associated with a human disease, preferably a NOVX-associated
XX disorder, or for treating or preventing a NOVX-associated disorder in a
XX subject, preferably human. The isolated protein, its encoding
XX polynucleotide or an antibody created from the protein are also useful
XX for treating or preventing metabolic disorders, diabetes, obesity,
XX infectious disease, anorexia, neurodegenerative disorder, Alzheimer's
XX disease, Parkinson's disorder, immune disorders, hematopoietic
XX disorders, and various dyslipidaemias, metabolic disturbances associated
XX with obesity, the metabolic syndrome X, wasting disorders associated with

Search completed: July 1, 2004, 13:46:36
Job time : 54 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2004, 13:48:47 ; Search time 44.5 Seconds
(without alignments)
2993.921 Million cell updates/sec

Title: US-10-091-085-3
Perfect score: 2250
Sequence: 1 MATSGTFFMLVSCVCSA.....ETGVALGATFHLLOSGLISH 428

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1276540 seqs, 311283816 residues 1276540
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications AA.*
- 1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 7: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
 - 10: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
 - 11: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 12: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 13: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 14: /cgn2_6/prodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 15: /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 16: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 17: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep.*
 - 18: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2250	100.0	428	13	US-10-091-085-3
2	2250	100.0	428	13	US-10-091-085-5
3	2250	100.0	428	13	US-10-092-063-3
4	2250	100.0	428	13	US-10-092-063-5
5	2250	100.0	428	14	US-10-286-926-3
6	2250	100.0	428	15	US-10-286-926-5
7	2250	100.0	428	15	US-10-231-913-127
8	2236	99.4	428	16	US-10-408-765A-2296
9	2235	99.3	428	13	US-10-091-085-7
10	2235	99.3	428	13	US-10-092-063-7
11	2235	99.3	428	14	US-10-286-926-7
12	2104	93.5	405	13	US-10-092-063-25
13	2104	93.5	405	14	US-10-286-926-25
14	1994.5	88.6	427	15	US-10-231-913-126
15	1837.5	81.7	465	13	US-10-092-063-39

16	1660	73.8	330	9	US-09-925-299-876	Sequence 876, App
17	1660	73.8	330	10	US-09-925-299-876	Sequence 876, App
18	999	44.4	484	13	US-10-092-063-27	Sequence 27, Appl
19	999	44.4	484	14	US-10-286-926-27	Sequence 27, Appl
20	999	44.4	484	15	US-10-231-913-123	Sequence 123, Appl
21	996	44.3	467	15	US-10-231-913-36	Sequence 36, Appl
22	993.5	44.2	379	15	US-10-231-913-271	Sequence 271, App
23	992	44.1	484	15	US-10-231-913-124	Sequence 124, App
24	989	44.0	446	15	US-10-231-913-38	Sequence 38, Appl
25	986	43.8	455	15	US-10-231-913-125	Sequence 125, App
26	616.5	27.4	479	15	US-10-369-493-6447	Sequence 6447, Ap
27	590	26.2	476	9	US-09-835-147-4	Sequence 4, Appli
28	538.5	23.9	556	15	US-10-369-493-2169	Sequence 2169, Ap
29	533	23.7	476	9	US-09-835-147-3	Sequence 3, Appli
30	507.5	22.6	459	9	US-09-129-112-9	Sequence 9, Appli
31	503.5	22.4	467	9	US-09-129-112-19	Sequence 19, Appl
32	503	22.4	410	15	US-10-231-913-272	Sequence 272, App
33	503	22.4	467	12	US-10-425-114-45875	Sequence 45875, A
34	498	22.1	462	9	US-09-139-112-2	Sequence 2, Appli
35	498	22.1	534	16	US-10-437-963-132351	Sequence 132351,
36	491.5	21.8	462	9	US-09-129-112-15	Sequence 15, Appl
37	491	21.8	433	12	US-10-425-114-36762	Sequence 36762, A
38	490	21.8	467	12	US-10-424-599-230158	Sequence 230158,
39	483.5	21.5	518	15	US-10-369-493-1713	Sequence 1713, Ap
40	474.5	21.1	443	12	US-10-425-114-49933	Sequence 49933, A
41	459.5	20.4	457	12	US-10-425-114-51762	Sequence 51762, A
42	459.5	20.4	459	12	US-10-424-599-145076	Sequence 145076,
43	455.5	20.2	472	16	US-10-437-963-151710	Sequence 151710,
44	445.5	19.8	465	12	US-10-425-114-37241	Sequence 37241, A
45	445.5	19.8	465	12	US-10-425-114-49932	Sequence 49932, A

ALIGNMENTS

RESULT 1

US-10-091-085-3
; Sequence 3, Application US/10091085
; Publication No. US2002014672A1
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 28110/35761
; CURRENT APPLICATION NUMBER: US/10/091,085
; CURRENT FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-091-085-3

Query Match 100.0%; Score 2250; DB 13; Length 428;
Best Local Similarity 100.0%; Pred. NO. 5.6e-215;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MATSGTFFMLVSCVCSA	VHRRNQTFEGIFLSSMCPINVSA	SLYGINFDAQSTGT	60
DB	1	MATSGTFFMLVSCVCSA	VHRRNQTFEGIFLSSMCPINVSA	SLYGINFDAQSTGT	60

QY 61 RIHYVTFVQKMPQQLPILGEVFDVSKPGLSAFVDPQKQAGTVOGLLEVAKDSIPRSHW 120
 DB 61 RIHYVTFVQKMPQQLPILGEVFDVSKPGLSAFVDPQKQAGTVOGLLEVAKDSIPRSHW 120
 QY 121 KKTPTVVLKATAGLRLLPEHAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWTV 180
 DB 121 KKTPTVVLKATAGLRLLPEHAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWTV 180
 QY 181 NFLTGOLHGHROETVGTDLGGASTQITFLPQFECTLEQTPRGYLTSTSEMFNSTYKLYTH 240
 DB 181 NFLTGOLHGHROETVGTDLGGASTQITFLPQFECTLEQTPRGYLTSTSEMFNSTYKLYTH 240
 QY 241 SYLGFGGLKAARLALATLGALETGTGHTFRSACLPRWLEAEWIFGGVKKYQYGGNQGGEVGF 300
 DB 241 SYLGFGGLKAARLALATLGALETGTGHTFRSACLPRWLEAEWIFGGVKKYQYGGNQGGEVGF 300
 QY 301 EPCYAEVLRVVRGKLHQPVEVQSGSYAFSYYYDRAVDTMDIDYKGGILKVEDPERKAR 360
 DB 301 EPCYAEVLRVVRGKLHQPVEVQSGSYAFSYYYDRAVDTMDIDYKGGILKVEDPERKAR 360
 QY 361 EVCNLENTSSGPFLLCMDSYITALLKDGFGFADSTVLQLTKKVNNIETGWLGAATPHL 420
 DB 361 EVCNLENTSSGPFLLCMDSYITALLKDGFGFADSTVLQLTKKVNNIETGWLGAATPHL 420
 QY 421 LQSLGISH 428
 DB 421 LQSLGISH 428

RESULT 2

US-10-091-085-5
 ; Sequence 5, Application US/10091085
 ; Publication No. US20020146772A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ford, John
 ; APPLICANT: Mulero, Julio
 ; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
 ; FILE REFERENCE: 28110/35761
 ; CURRENT APPLICATION NUMBER: US/10/091,085
 ; CURRENT FILING DATE: 2002-03-05
 ; PRIOR APPLICATION NUMBER: 09/350,836
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR APPLICATION NUMBER: 09/273,447
 ; PRIOR FILING DATE: 1999-03-19
 ; PRIOR APPLICATION NUMBER: 09/118,205
 ; PRIOR FILING DATE: 1998-07-16
 ; PRIOR APPLICATION NUMBER: 09/122,449
 ; PRIOR FILING DATE: 1998-07-24
 ; PRIOR APPLICATION NUMBER: 09/244,444
 ; PRIOR FILING DATE: 1999-02-04
 ; NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 428
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; ORGANISM: Homo sapiens

Query Match 100.0%; Score 2250; DB 13; Length 428;
 Best Local Similarity 100.0%; Pred. No. 5.6e-215;
 Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATSNGTVPFMLVWSCVSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
 DB 1 MATSNGTVPFMLVWSCVSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
 QY 61 RIHYVTFVQKMPQQLPILGEVFDVSKPGLSAFVDPQKQAGTVOGLLEVAKDSIPRSHW 120
 DB 61 RIHYVTFVQKMPQQLPILGEVFDVSKPGLSAFVDPQKQAGTVOGLLEVAKDSIPRSHW 120
 QY 121 KKTPTVVLKATAGLRLLPEHAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWTV 180

DB 121 KKTPTVVLKATAGLRLLPEHAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWTV 180
 QY 181 NFLTGOLHGHROETVGTDLGGASTQITFLPQFECTLEQTPRGYLTSTSEMFNSTYKLYTH 240
 DB 181 NFLTGOLHGHROETVGTDLGGASTQITFLPQFECTLEQTPRGYLTSTSEMFNSTYKLYTH 240
 QY 241 SYLGFGGLKAARLALATLGALETGTGHTFRSACLPRWLEAEWIFGGVKKYQYGGNQGGEVGF 300
 DB 241 SYLGFGGLKAARLALATLGALETGTGHTFRSACLPRWLEAEWIFGGVKKYQYGGNQGGEVGF 300
 QY 301 EPCYAEVLRVVRGKLHQPVEVQSGSYAFSYYYDRAVDTMDIDYKGGILKVEDPERKAR 360
 DB 301 EPCYAEVLRVVRGKLHQPVEVQSGSYAFSYYYDRAVDTMDIDYKGGILKVEDPERKAR 360
 QY 361 EVCNLENTSSGPFLLCMDSYITALLKDGFGFADSTVLQLTKKVNNIETGWLGAATPHL 420
 DB 361 EVCNLENTSSGPFLLCMDSYITALLKDGFGFADSTVLQLTKKVNNIETGWLGAATPHL 420
 QY 421 LQSLGISH 428
 DB 421 LQSLGISH 428

RESULT 3

US-10-092-063-3
 ; Sequence 3, Application US/10092063
 ; Publication No. US20020173005A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ford, John
 ; APPLICANT: Mulero, Julio
 ; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDE
 ; FILE REFERENCE: 28110/35908
 ; CURRENT APPLICATION NUMBER: US/10/092,063
 ; CURRENT FILING DATE: 2002-03-05
 ; PRIOR APPLICATION NUMBER: 09/370,265
 ; PRIOR FILING DATE: 2002-01-31
 ; PRIOR APPLICATION NUMBER: PCT/US99/16180
 ; PRIOR FILING DATE: 1999-07-16
 ; PRIOR APPLICATION NUMBER: 09/350,836
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR APPLICATION NUMBER: 09/273,447
 ; PRIOR FILING DATE: 1999-03-19
 ; PRIOR APPLICATION NUMBER: 09/244,444
 ; PRIOR FILING DATE: 1999-02-04
 ; PRIOR APPLICATION NUMBER: 09/122,449
 ; PRIOR FILING DATE: 1998-07-24
 ; PRIOR APPLICATION NUMBER: 09/118,205
 ; PRIOR FILING DATE: 1998-07-16
 ; NUMBER OF SEQ ID NOS: 39
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 428
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; ORGANISM: Homo sapiens

Query Match 100.0%; Score 2250; DB 13; Length 428;
 Best Local Similarity 100.0%; Pred. No. 5.6e-215;
 Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATSNGTVPFMLVWSCVSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
 DB 1 MATSNGTVPFMLVWSCVSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
 QY 61 RIHYVTFVQKMPQQLPILGEVFDVSKPGLSAFVDPQKQAGTVOGLLEVAKDSIPRSHW 120
 DB 61 RIHYVTFVQKMPQQLPILGEVFDVSKPGLSAFVDPQKQAGTVOGLLEVAKDSIPRSHW 120
 QY 121 KKTPTVVLKATAGLRLLPEHAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWTV 180
 DB 121 KKTPTVVLKATAGLRLLPEHAKALLFEVKEIFRKSPFLVPKGSVSIMDGSDEGILAWTV 180
 QY 181 NFLTGOLHGHROETVGTDLGGASTQITFLPQFECTLEQTPRGYLTSTSEMFNSTYKLYTH 240

Db 181 NLTGQLHGRQETVGTLDLGASTQITFLPQFEXTLEQTPRGYLTSFEMFNSTYKLYTH 240
QY 241 SYLGFLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYCGNQEYGF 300
Db 241 SYLGFLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYCGNQEYGF 300
QY 301 EPCYAEVLVRVGRKLHQPVEVQSGFYAFSYYYDRAVDMDIDYKGGILKVEDPERKAR 360
Db 301 EPCYAEVLVRVGRKLHQPVEVQSGFYAFSYYYDRAVDMDIDYKGGILKVEDPERKAR 360
QY 361 EYCDNLENTSGSPFLCMLDLSYITALLKDGFGFADSTVLQTLTKKNNIETGHALGATPHL 420
Db 361 EYCDNLENTSGSPFLCMLDLSYITALLKDGFGFADSTVLQTLTKKNNIETGHALGATPHL 420
QY 421 LQSLGISH 428
Db 421 LQSLGISH 428

RESULT 4
US-10-092-063-5
; Sequence 5, Application US/10092063
; Publication No. US20020173005A1
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES
; FILE REFERENCE: 28110/35908
; CURRENT APPLICATION NUMBER: US/10/092,063
; CURRENT FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-063-5

Query Match 100.0%; Score 2250; DB 13; Length 428;
Best Local Similarity 100.0%; Pred. No. 5.6e-215;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATSWGTVFFMLVVCVCSAVSHRNQOTWEEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSWGTVFFMLVVCVCSAVSHRNQOTWEEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
QY 61 RIHVYTFVQKMPGQLPILEGVEFDSVAPGLSAFVDPQKQGAETVQGLLEVAKDSIPRSHW 120
Db 61 RIHVYTFVQKMPGQLPILEGVEFDSVAPGLSAFVDPQKQGAETVQGLLEVAKDSIPRSHW 120
QY 121 KCTPVLKATAGLLPEHAKALLFEVKIEPRKSPFLVPKGSVIMDGSDEGILLAWTV 180
Db 121 KCTPVLKATAGLLPEHAKALLFEVKIEPRKSPFLVPKGSVIMDGSDEGILLAWTV 180
QY 181 NLTGQLHGRQETVGTLDLGASTQITFLPQFEXTLEQTPRGYLTSFEMFNSTYKLYTH 240
Db 181 NLTGQLHGRQETVGTLDLGASTQITFLPQFEXTLEQTPRGYLTSFEMFNSTYKLYTH 240

QY 241 SYLGFLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYCGNQEYGF 300
Db 241 SYLGFLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYCGNQEYGF 300
QY 301 EPCYAEVLVRVGRKLHQPVEVQSGFYAFSYYYDRAVDMDIDYKGGILKVEDPERKAR 360
Db 301 EPCYAEVLVRVGRKLHQPVEVQSGFYAFSYYYDRAVDMDIDYKGGILKVEDPERKAR 360
QY 361 EYCDNLENTSGSPFLCMLDLSYITALLKDGFGFADSTVLQTLTKKNNIETGHALGATPHL 420
Db 361 EYCDNLENTSGSPFLCMLDLSYITALLKDGFGFADSTVLQTLTKKNNIETGHALGATPHL 420
QY 421 LQSLGISH 428
Db 421 LQSLGISH 428

RESULT 5
US-10-286-926-3
; Sequence 3, Application US/10286926
; Publication No. US20030175752A1
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
; FILE REFERENCE: 28110/36457CON
; CURRENT APPLICATION NUMBER: US/10/286,926
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/122449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-286-926-3

Query Match 100.0%; Score 2250; DB 14; Length 428;
Best Local Similarity 100.0%; Pred. No. 5.6e-215;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATSWGTVFFMLVVCVCSAVSHRNQOTWEEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSWGTVFFMLVVCVCSAVSHRNQOTWEEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
QY 61 RIHVYTFVQKMPGQLPILEGVEFDSVAPGLSAFVDPQKQGAETVQGLLEVAKDSIPRSHW 120
Db 61 RIHVYTFVQKMPGQLPILEGVEFDSVAPGLSAFVDPQKQGAETVQGLLEVAKDSIPRSHW 120
QY 121 KCTPVLKATAGLLPEHAKALLFEVKIEPRKSPFLVPKGSVIMDGSDEGILLAWTV 180
Db 121 KCTPVLKATAGLLPEHAKALLFEVKIEPRKSPFLVPKGSVIMDGSDEGILLAWTV 180
QY 181 NLTGQLHGRQETVGTLDLGASTQITFLPQFEXTLEQTPRGYLTSFEMFNSTYKLYTH 240

Db 191 NLTGQLHGRQETVGTLDLGGASTQITFLPQPEKTLBOPTPRGYLTSPFENSTYKLYTH 240
Qy 241 SYLAFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVYQYGGNQEVEVGF 300
Db 241 SYLAFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVYQYGGNQEVEVGF 300
Qy 301 EPCYAEVLVRVGRKLGHPBEVQSGSFYAFSYYYDRAVDMDIDYKGGILKVEDPFRKAR 360
Db 301 EPCYAEVLVRVGRKLGHPBEVQSGSFYAFSYYYDRAVDMDIDYKGGILKVEDPFRKAR 360
Qy 361 EVCNLENFTSGSPFLCMLDSYITALLXDGFGFADSTVLQLTKKVNNIETGVALGATFHL 420
Db 361 EVCNLENFTSGSPFLCMLDSYITALLXDGFGFADSTVLQLTKKVNNIETGVALGATFHL 420
Qy 421 LQSLGISH 428
Db 421 LQSLGISH 428
RESULT 6
US-10-286-926-5
; Sequence 5, Application US/10286926
; Publication No. US20030175752A1
; GENERAL INFORMATION:
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 28110/36457CON
; CURRENT APPLICATION NUMBER: US/10/286,926
; PRIOR FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/122449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/24444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-286-926-5
Query Match 100.0%; Score 2250; DB 14; Length 428;
Best Local Similarity 100.0%; Pred. No. 5.6e-215;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MATSGVTFFVLMVSCVSAVSHRNQQTWFFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSGVTFFVLMVSCVSAVSHRNQQTWFFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Qy 61 RIHYVTFVQKMPGQLPILEGVFDVSKPGLSAFVDQPKQGAETVQGLLEAKDSIPRSHW 120
Db 61 RIHYVTFVQKMPGQLPILEGVFDVSKPGLSAFVDQPKQGAETVQGLLEAKDSIPRSHW 120
Qy 121 KKTTPVWLKATAGLLELPEHKAKALLFEVKEIFRKSPLVPKGSVINDSGDEGLAWTV 180

Db 121 KKTTPVWLKATAGLLELPEHKAKALLFEVKEIFRKSPLVPKGSVINDSGDEGLAWTV 180
Qy 181 NLTGQLHGRQETVGTLDLGGASTQITFLPQPEKTLBOPTPRGYLTSPFENSTYKLYTH 240
Db 181 NLTGQLHGRQETVGTLDLGGASTQITFLPQPEKTLBOPTPRGYLTSPFENSTYKLYTH 240
Qy 241 SYLAFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVYQYGGNQEVEVGF 300
Db 241 SYLAFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVYQYGGNQEVEVGF 300
Qy 301 EPCYAEVLVRVGRKLGHPBEVQSGSFYAFSYYYDRAVDMDIDYKGGILKVEDPFRKAR 360
Db 301 EPCYAEVLVRVGRKLGHPBEVQSGSFYAFSYYYDRAVDMDIDYKGGILKVEDPFRKAR 360
Qy 361 EVCNLENFTSGSPFLCMLDSYITALLXDGFGFADSTVLQLTKKVNNIETGVALGATFHL 420
Db 361 EVCNLENFTSGSPFLCMLDSYITALLXDGFGFADSTVLQLTKKVNNIETGVALGATFHL 420
Qy 421 LQSLGISH 428
Db 421 LQSLGISH 428
RESULT 7
US-10-231-913-127
; Sequence 127, Application US/10231913
; Publication No. US20040005576A1
; GENERAL INFORMATION:
; APPLICANT: Guo, Xiaojia S.
; APPLICANT: Li, Li
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Vernet, Corine A.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Edinger, Schiomit
; APPLICANT: Peyman, John A.
; APPLICANT: Stone, David J.
; APPLICANT: Billeman, Karen
; APPLICANT: Gangolli, Baha A.
; APPLICANT: Boldog, Ference L.
; APPLICANT: Colman, Steven D.
; APPLICANT: Eisen, Andrew J.
; APPLICANT: Liu, Xiachong
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Zerhusen, Bryan D.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-216
; CURRENT APPLICATION NUMBER: US/10/231,913
; CURRENT FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: 60/251,660
; PRIOR FILING DATE: 2000-12-06
; PRIOR APPLICATION NUMBER: 60/255,029
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/260,326
; PRIOR FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 60/263,800
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/269,942
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/286,183
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 60/313,627
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/318,712
; PRIOR FILING DATE: 2001-09-12
; NUMBER OF SEQ ID NOS: 292
; SOFTWARE: PatentIn Ver. 2.1

```

; SEQ ID NO 127
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-231-913-127

Query Match      100.0%; Score 2250; DB 15; Length 428;
Best Local Similarity 100.0%; Pred. No. 5.6e-215;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATSWGTVFVFLVSVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
DB |||||
QY 61 RIHVYTFVQKMPGQLPILLEGVDSVKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120
DB |||||
QY 61 RIHVYTFVQKMPGQLPILLEGVDSVKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120
DB |||||
QY 121 KXTPVVLKATAGRLRLPEHAKALLFEVKIIFRKSPPFLVPKGSVIMDGSDEGILAWTV 180
DB |||||
QY 181 NFLTGQHGHRQETVGTLDLGGASTQITFLPQFEKTLQTPRGYLTSEMFNSTYKLYTH 240
DB |||||
QY 181 NFLTGQHGHRQETVGTLDLGGASTQITFLPQFEKTLQTPRGYLTSEMFNSTYKLYTH 240
DB |||||
QY 241 SYLGFGGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGF 300
DB |||||
QY 301 EPCYAEVLVRVGRKGLHQPVEVQSGFYAFSYVDRAVDTMDIDYEKGGILKVEDPERKAR 360
DB |||||
QY 301 EPCYAEVLVRVGRKGLHQPVEVQSGFYAFSYVDRAVDTMDIDYEKGGILKVEDPERKAR 360
DB |||||
QY 361 EVCNLENFTSGSPFLCWDLSYITALLKDGFGFADSTVLQLTKKVNNIETGVALGATFHL 420
DB |||||
QY 421 LQSLGISH 428
DB |||||
QY 421 LQSLGISH 428
DB |||||

RESULT 9
US-10-091-085-7
; Sequence 7, Application US/10091085
; Publication No. US20020146772A1
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 28110/35761
; CURRENT APPLICATION NUMBER: US/10/091,085
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-091-085-7

Query Match      99.3%; Score 2235; DB 13; Length 428;
Best Local Similarity 99.3%; Pred. No. 1.7e-213;
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MATSWGTVFVFLVSVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
DB |||||
QY 61 RIHVYTFVQKMPGQLPILLEGVDSVKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120
DB |||||
QY 61 RIHVYTFVQKMPGQLPILLEGVDSVKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120
DB |||||

; SEQ ID NO 127
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-231-913-127

Query Match      100.0%; Score 2250; DB 15; Length 428;
Best Local Similarity 100.0%; Pred. No. 5.6e-215;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATSWGTVFVFLVSVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
DB |||||
QY 61 RIHVYTFVQKMPGQLPILLEGVDSVKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120
DB |||||
QY 61 RIHVYTFVQKMPGQLPILLEGVDSVKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120
DB |||||
QY 121 KXTPVVLKATAGRLRLPEHAKALLFEVKIIFRKSPPFLVPKGSVIMDGSDEGILAWTV 180
DB |||||
QY 181 NFLTGQHGHRQETVGTLDLGGASTQITFLPQFEKTLQTPRGYLTSEMFNSTYKLYTH 240
DB |||||
QY 181 NFLTGQHGHRQETVGTLDLGGASTQITFLPQFEKTLQTPRGYLTSEMFNSTYKLYTH 240
DB |||||
QY 241 SYLGFGGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGF 300
DB |||||
QY 241 SYLGFGGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGF 300
DB |||||
QY 301 EPCYAEVLVRVGRKGLHQPVEVQSGFYAFSYVDRAVDTMDIDYEKGGILKVEDPERKAR 360
DB |||||
QY 301 EPCYAEVLVRVGRKGLHQPVEVQSGFYAFSYVDRAVDTMDIDYEKGGILKVEDPERKAR 360
DB |||||
QY 361 EVCNLENFTSGSPFLCWDLSYITALLKDGFGFADSTVLQLTKKVNNIETGVALGATFHL 420
DB |||||
QY 421 LQSLGISH 428
DB |||||
QY 421 LQSLGISH 428
DB |||||

RESULT 8
US-10-408-765A-2296
; Sequence 2296, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2296
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-408-765A-2296

Query Match      99.4%; Score 2236; DB 16; Length 428;
Best Local Similarity 99.5%; Pred. No. 1.4e-213;
Matches 426; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MATSWGTVFVFLVSVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
```

QY 121 KKT PVLKATAGLRLPEHKAALLFEVKEIFRKSFLVPKGSVIMDSGDEGILAWTV 180
DB 121 KKT PVLKATAGLRLPEHKAALLFEVKEIFRKSFLVPKGSVIMDSGDEGILAWTV 180
QY 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSPFEMFNSTYKLYTH 240
DB 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSPFEMFNSTYKLYTH 240
QY 241 SYLGFGKAAARLATIAGLETGTGHTFRSACLPRWLEAEWIFGVKYQVGNQGEVGF 300
DB 241 SYLGFGKAAARLATIAGLETGTGHTFRSACLPRWLEAEWIFGVKYQVGNQGEVGF 300
QY 301 EPCYAEVLVRVVRGKLGHPVEVQSGSFYAFSYYYDRAVDMDIDYKGGILKVEDFERKAR 360
DB 301 EPCYAEVLVRVVRGKLGHPVEVQSGSFYAFSYYYDRAVDMDIDYKGGILKVEDFERKAR 360
QY 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGPADSTVLQTKKVNNIETGWAIGATFHL 420
DB 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGPADSTVLQTKKVNNIETGWAIGATFHL 420
QY 421 LQSLGISH 428
DB 421 LQSLGISH 428

RESULT 10

US-10-092-063-7
; Sequence 7, Application US/10092063
; Publication No. US20020173005A1
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES
; FILE REFERENCE: 28110/35908
; CURRENT APPLICATION NUMBER: US/10/092,063
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 7
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-063-7

Query Match 99.3%; Score 2235; DB 13; Length 428;
Best Local Similarity 99.3%; Pred. No. 1.7e-213;
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MATSWGTVFFMLVVCVCSAVSHRNQQTWPEGIFLSSMCPINVSASTLYGIMPDAGSTGT 60
DB 1 MATSWGTVFFMLVVCVCSAVSHRNQQTWPEGIFLSSMCPINVSASTLYGIMPDAGSTGT 60
QY 61 RIHYVTFVQKMPGQLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
DB 61 RIHYVTFVQKMPGQLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
QY 121 KKT PVLKATAGLRLPEHKAALLFEVKEIFRKSFLVPKGSVIMDSGDEGILAWTV 180

DB 121 KKT PVLKATAGLRLPEHKAALLFEVKEIFRKSFLVPKGSVIMDSGDEGILAWTV 180
QY 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSPFEMFNSTYKLYTH 240
DB 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSPFEMFNSTYKLYTH 240
QY 241 SYLGFGKAAARLATIAGLETGTGHTFRSACLPRWLEAEWIFGVKYQVGNQGEVGF 300
DB 241 SYLGFGKAAARLATIAGLETGTGHTFRSACLPRWLEAEWIFGVKYQVGNQGEVGF 300
QY 301 EPCYAEVLVRVVRGKLGHPVEVQSGSFYAFSYYYDRAVDMDIDYKGGILKVEDFERKAR 360
DB 301 EPCYAEVLVRVVRGKLGHPVEVQSGSFYAFSYYYDRAVDMDIDYKGGILKVEDFERKAR 360
QY 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGPADSTVLQTKKVNNIETGWAIGATFHL 420
DB 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGPADSTVLQTKKVNNIETGWAIGATFHL 420
QY 421 LQSLGISH 428
DB 421 LQSLGISH 428

RESULT 11

US-10-286-926-7
; Sequence 7, Application US/10286926
; Publication No. US20030175752A1
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
; FILE REFERENCE: 28110/36457CON
; CURRENT APPLICATION NUMBER: US/10/286,926
; PRIOR FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/122449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 7
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-286-926-7

Query Match 99.3%; Score 2235; DB 14; Length 428;
Best Local Similarity 99.3%; Pred. No. 1.7e-213;
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MATSWGTVFFMLVVCVCSAVSHRNQQTWPEGIFLSSMCPINVSASTLYGIMPDAGSTGT 60
DB 1 MATSWGTVFFMLVVCVCSAVSHRNQQTWPEGIFLSSMCPINVSASTLYGIMPDAGSTGT 60
QY 61 RIHYVTFVQKMPGQLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
DB 61 RIHYVTFVQKMPGQLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

QY 121 KKTPTVVLKATAGRLLLPEHAKALLFEVKEIFRKSPLVPKGSVIMDGSDEGILLAWTV 180
DB 121 KKTPTVVLKATAGRLLLPEHAKALLFEVKEIFRKSPLVPKGSVIMDGSDEGILLAWTV 180
QY 181 NFLTQGLHGRQETVGTLDLGGASTQITFLPQFEKLEQTPRGYLTSEFMPNSTYKLYTH 240
DB 181 NFLTQGLHGRQETVGTLDLGGASTQITFLPQFEKLEQTPRGYLTSEFMPNSTYKLYTH 240
QY 241 SYLGFGKKAARLATLGALETGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNOGEVGF 300
DB 241 SYLGFGKKAARLATLGALETGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNOGEVGF 300
QY 301 EPCYAEVLRVVRGKHLQHEPVEVQSGFYAFSYDDRAVDTMDIDYKGGILKVEDFERKAR 360
DB 301 EPCYAEVLRVVRGKHLQHEPVEVQSGFYAFSYDDRAVDTMDIDYKGGILKVEDFERKAR 360
QY 361 EVCNLENFTSGSPFLCMLSYITALLKDGFGFADSTVLQ 400
DB 361 EVCNLENFTSGSPFLCMLSYITALLKDGFGFADSTVLQ 400
QY 421 LQSLGISH 428
DB 421 LQSLGISH 428
RESULT 12
US-10-092-063-25
; Sequence 25, Application US/10092063
; Publication No. US20020173005A1
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES
; FILE REFERENCE: 28110/35908
; CURRENT APPLICATION NUMBER: US/10/092.063
; CURRENT FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-063-25

Query Match 93.5%; Score 2104; DB 13; Length 405;
Best Local Similarity 100.0%; Pred. No. 1.8e-200;
Matches 400; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATSWGTVPFVFLVSVCSAVSHRNQQTWFEGLFSMCPINVSASTLYGIMFDAGSTGT 60
DB 1 MATSWGTVPFVFLVSVCSAVSHRNQQTWFEGLFSMCPINVSASTLYGIMFDAGSTGT 60
QY 61 RIHYVTFVQKMPGQLPILGEVDFSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
DB 61 RIHYVTFVQKMPGQLPILGEVDFSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
QY 121 KKTPTVVLKATAGRLLLPEHAKALLFEVKEIFRKSPLVPKGSVIMDGSDEGILLAWTV 180

DB 121 KKTPTVVLKATAGRLLLPEHAKALLFEVKEIFRKSPLVPKGSVIMDGSDEGILLAWTV 180
QY 181 NFLTQGLHGRQETVGTLDLGGASTQITFLPQFEKLEQTPRGYLTSEFMPNSTYKLYTH 240
DB 181 NFLTQGLHGRQETVGTLDLGGASTQITFLPQFEKLEQTPRGYLTSEFMPNSTYKLYTH 240
QY 241 SYLGFGKKAARLATLGALETGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNOGEVGF 300
DB 241 SYLGFGKKAARLATLGALETGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNOGEVGF 300
QY 301 EPCYAEVLRVVRGKHLQHEPVEVQSGFYAFSYDDRAVDTMDIDYKGGILKVEDFERKAR 360
DB 301 EPCYAEVLRVVRGKHLQHEPVEVQSGFYAFSYDDRAVDTMDIDYKGGILKVEDFERKAR 360
QY 361 EVCNLENFTSGSPFLCMLSYITALLKDGFGFADSTVLQ 400
DB 361 EVCNLENFTSGSPFLCMLSYITALLKDGFGFADSTVLQ 400
RESULT 13
US-10-286-926-25
; Sequence 25, Application US/10286926
; Publication No. US20030175752A1
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 28110/36457CON
; CURRENT APPLICATION NUMBER: US/10/286,926
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-286-926-25

Query Match 93.5%; Score 2104; DB 14; Length 405;
Best Local Similarity 100.0%; Pred. No. 1.8e-200;
Matches 400; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATSWGTVPFVFLVSVCSAVSHRNQQTWFEGLFSMCPINVSASTLYGIMFDAGSTGT 60
DB 1 MATSWGTVPFVFLVSVCSAVSHRNQQTWFEGLFSMCPINVSASTLYGIMFDAGSTGT 60
QY 61 RIHYVTFVQKMPGQLPILGEVDFSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
DB 61 RIHYVTFVQKMPGQLPILGEVDFSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
QY 121 KKTPTVVLKATAGRLLLPEHAKALLFEVKEIFRKSPLVPKGSVIMDGSDEGILLAWTV 180
DB 121 KKTPTVVLKATAGRLLLPEHAKALLFEVKEIFRKSPLVPKGSVIMDGSDEGILLAWTV 180

QY 181 NFLTQGLHGRQETVGTLDLGGASTQITFLPQFENKLEQTPRGYLTSPFENSTYKLYTH 240
 DB 181 NFLTQGLHGRQETVGTLDLGGASTQITFLPQFENKLEQTPRGYLTSPFENSTYKLYTH 240
 QY 241 SYLGFGKKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQGEGVGF 300
 DB 241 SYLGFGKKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQGEGVGF 300
 QY 301 EPCVAEVLVRVGRKLGHPQEEVQSGFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
 DB 301 EPCVAEVLVRVGRKLGHPQEEVQSGFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
 QY 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQ 400
 DB 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQ 400

RESULT 14
 US-10-231-913-126
 ; Sequence 126, Application US/10231913
 ; Publication No. US20040005576A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Guo, Xiaojia S.
 ; APPLICANT: Li, Li
 ; APPLICANT: Patturajan, Meera
 ; APPLICANT: Shinkets, Richard A.
 ; APPLICANT: Casman, Stacie J.
 ; APPLICANT: Malyankar, Uriel M.
 ; APPLICANT: Tchernev, Velizar T.
 ; APPLICANT: Vernet, Corine A.
 ; APPLICANT: Spytek, Kimberly A.
 ; APPLICANT: Shenoy, Suresh G.
 ; APPLICANT: Alsobrook II, John P.
 ; APPLICANT: Edinger, Schlomit
 ; APPLICANT: Peyman, John A.
 ; APPLICANT: Stone, David J.
 ; APPLICANT: Ellerman, Karen
 ; APPLICANT: Gangolli, Esha A.
 ; APPLICANT: Boldog, Ference L.
 ; APPLICANT: Colman, Steven D.
 ; APPLICANT: Eisen, Andrew J.
 ; APPLICANT: Liu, Xiaohong
 ; APPLICANT: Padigaru, Muralidhara
 ; APPLICANT: Spaderna, Steven K.
 ; APPLICANT: Zerhusen, Bryan D.
 ; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 21402-216
 ; CURRENT APPLICATION NUMBER: US/10/231,913
 ; CURRENT FILING DATE: 2002-08-30
 ; PRIOR APPLICATION NUMBER: 60/251,660
 ; PRIOR FILING DATE: 2000-12-06
 ; PRIOR APPLICATION NUMBER: 60/255,029
 ; PRIOR FILING DATE: 2000-12-12
 ; PRIOR APPLICATION NUMBER: 60/260,326
 ; PRIOR FILING DATE: 2001-01-08
 ; PRIOR APPLICATION NUMBER: 60/263,800
 ; PRIOR FILING DATE: 2001-01-24
 ; PRIOR APPLICATION NUMBER: 60/269,942
 ; PRIOR FILING DATE: 2001-02-20
 ; PRIOR APPLICATION NUMBER: 60/286,183
 ; PRIOR FILING DATE: 2001-04-24
 ; PRIOR APPLICATION NUMBER: 60/313,627
 ; PRIOR FILING DATE: 2001-08-20
 ; PRIOR APPLICATION NUMBER: 60/318,712
 ; PRIOR FILING DATE: 2001-09-12
 ; NUMBER OF SEQ ID NOS: 292
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 126
 ; TYPE: PRT
 ; LENGTH: 427
 ; ORGANISM: Mus musculus
 US-10-231-913-126

Query Match 81.7%; Score 1837.5; DB 13; Length 465;
 Best Local Similarity 84.0%; Pred. No. 7.8e-174;

Query Match 88.6%; Score 1994.5; DB 15; Length 427;
 Best Local Similarity 88.3%; Pred. No. 1.6e-189;
 Matches 377; Conservative 24; Mismatches 25; Indels 1; Gaps 1;
 QY 1 MATSWGVVPMVLVVCVSAVSHRNQOTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
 DB 1 MATSWGVVPMVLVVCVSAVSHRNQOTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 59
 QY 61 RIHYTVPVQKQPLILEGEVFDTSVKPGLSAFVDPQKGAETVQGLLEVAKDSIPRSHW 120
 DB 60 RIHYTVPVQKQPLILEGEVFDTSVKPGLSAFVDPQKGAETVQGLLEVAKDSIPRSHW 119
 QY 121 KKTVPVLKATAGRLLEPEHAKALLPEVKEIFPKSPFLVPKGSVIMDGSDEGILAWTV 180
 DB 120 ERTVPVLKATAGRLLEPEHAKALLPEVKEIFPKSPFLVPKGSVIMDGSDEGILAWTV 179
 QY 181 NFLTQGLHGRQETVGTLDLGGASTQITFLPQFENKLEQTPRGYLTSPFENSTYKLYTH 240
 DB 180 NFLTQGLHGRQETVGTLDLGGASTQITFLPQFENKLEQTPRGYLTSPFENSTYKLYTH 239
 QY 241 SYLGFGKKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQGEGVGF 300
 DB 240 SYLGFGKKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQGEGVGF 299
 QY 301 EPCVAEVLVRVGRKLGHPQEEVQSGFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
 DB 300 EPCVAEVLVRVGRKLGHPQEEVQSGFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 359
 QY 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQTKKVNNIETGALGATFHL 420
 DB 360 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQTKKVNNIETGALGATFHL 419
 QY 421 LQSLGIS 427
 DB 420 LQSLGIT 426

RESULT 15
 US-10-092-063-39
 ; Sequence 39, Application US/10092063
 ; Publication No. US20020173005A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ford, John
 ; APPLICANT: Mulero, Julio
 ; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDE
 ; FILE REFERENCE: 28110/35908
 ; CURRENT APPLICATION NUMBER: US/10/092,063
 ; CURRENT FILING DATE: 2002-03-05
 ; PRIOR APPLICATION NUMBER: 09/370,265
 ; PRIOR FILING DATE: 2002-01-31
 ; PRIOR APPLICATION NUMBER: PCT/US99/16180
 ; PRIOR FILING DATE: 1999-07-16
 ; PRIOR APPLICATION NUMBER: 09/350,836
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR APPLICATION NUMBER: 09/273,447
 ; PRIOR FILING DATE: 1999-03-19
 ; PRIOR APPLICATION NUMBER: 09/244,444
 ; PRIOR FILING DATE: 1999-02-04
 ; PRIOR APPLICATION NUMBER: 09/122,449
 ; PRIOR FILING DATE: 1998-07-24
 ; PRIOR APPLICATION NUMBER: 09/118,205
 ; PRIOR FILING DATE: 1998-07-16
 ; NUMBER OF SEQ ID NOS: 39
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 39
 ; TYPE: PRT
 ; LENGTH: 465
 ; ORGANISM: Mus musculus
 US-10-092-063-39

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2004, 13:44:42 ; Search time 16.5 Seconds
(without alignments)
1339.146 Million cell updates/sec

Title: US-10-091-085-7

Perfect score: 2252

Sequence: 1 MATSGTGVFFMLVSVCSA.....ETGWALGATPHLLQSLGISH 428

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTCUS.COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfilest.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2252	100.0	428	4	US-09-608-285A-7
2	2252	100.0	428	4	US-09-350-836B-7
3	2252	100.0	428	4	US-09-370-265-7
4	2252	100.0	428	4	US-09-557-800C-7
5	2252	100.0	428	4	US-09-370-625A-7
6	2235	99.2	428	4	US-09-608-285A-3
7	2235	99.2	428	4	US-09-608-285A-5
8	2235	99.2	428	4	US-09-240-639-6
9	2235	99.2	428	4	US-09-240-639-9
10	2235	99.2	428	4	US-09-350-836B-3
11	2235	99.2	428	4	US-09-350-836B-5
12	2235	99.2	428	4	US-09-370-265-3
13	2235	99.2	428	4	US-09-370-265-5
14	2235	99.2	428	4	US-09-557-800C-3
15	2235	99.2	428	4	US-09-557-800C-5
16	2235	99.2	428	4	US-09-370-625A-3
17	2235	99.2	428	4	US-09-370-625A-5
18	2089	92.8	405	4	US-09-608-285A-25
19	2089	92.8	405	4	US-09-370-265-25
20	2089	92.8	405	4	US-09-557-800C-25
21	2089	92.8	405	4	US-09-370-625A-25
22	1822.5	80.9	465	4	US-09-557-800C-56
23	1822.5	80.9	465	4	US-09-370-625A-39
24	1817.5	80.7	465	4	US-09-240-639-8
25	996	44.2	456	4	US-09-240-639-2
26	996	44.2	484	4	US-09-608-285A-27
27	996	44.2	484	4	US-09-370-265-27

28	996	44.2	484	4	US-09-557-800C-27	Sequence 27, Appl
29	996	44.2	484	4	US-09-370-625A-27	Sequence 27, Appl
30	816.5	36.3	471	4	US-09-608-285A-60	Sequence 60, Appl
31	507.5	22.3	467	4	US-09-129-112-19	Sequence 19, Appl
32	502.5	22.3	459	4	US-09-129-112-9	Sequence 9, Appl
33	493	21.9	462	4	US-09-129-112-2	Sequence 2, Appl
34	486.5	21.6	462	4	US-09-129-112-15	Sequence 15, Appl
35	484	21.5	455	4	US-09-240-639-10	Sequence 10, Appl
36	472.5	21.0	473	4	US-09-240-639-12	Sequence 12, Appl
37	441.5	19.6	454	4	US-09-240-639-11	Sequence 11, Appl
38	366	16.3	502	4	US-09-557-800C-55	Sequence 55, Appl
39	366	16.3	502	4	US-09-370-625A-38	Sequence 38, Appl
40	366	16.3	510	3	US-08-930-921-1	Sequence 1, Appl
41	357.5	15.9	529	4	US-09-240-639-4	Sequence 4, Appl
42	229	10.2	148	4	US-09-240-639-17	Sequence 17, Appl
43	201	8.9	153	4	US-09-240-639-13	Sequence 13, Appl
44	191	8.5	153	4	US-09-240-639-15	Sequence 15, Appl
45	189.5	8.4	150	4	US-09-240-639-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-09-608-285A-7
Sequence 7, Application US/09608285A
Patent No. 6335013
GENERAL INFORMATION:
APPLICANT: ~~Fora~~, John
APPLICANT: Mulero, Julio
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
TITLE OF INVENTION: POLYPEPTIDES
FILE REFERENCE: 28110/36570
CURRENT APPLICATION NUMBER: US/09/608,285A
CURRENT FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 09/583,231
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 09/557,800
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/481,238
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: 09/370,265
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: PCT/US99/16180
PRIOR FILING DATE: 1999-07-16
PRIOR APPLICATION NUMBER: 09/350,836
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 09/273,447
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 09/244,444
PRIOR FILING DATE: 1999-02-04
PRIOR APPLICATION NUMBER: 09/122,449
PRIOR FILING DATE: 1998-07-24
PRIOR APPLICATION NUMBER: 09/118,205
PRIOR FILING DATE: 1998-07-16
NUMBER OF SEQ ID NOS: 60
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 428
TYPE: PRT
ORGANISM: Homo sapiens
US-09-608-285A-7

Query Match 100.0%; Score 2252; DB 4; Length 428;
Best Local Similarity 100.0%; Pred No. 7, 8e-247; Indels 0; Gaps 0;
Matches 428; Conservative 0; Mismatches 0;

QY 1 MATSGTGVFFMLVSVCSAVSHRNQQTWTEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Db 1 MATSGTGVFFMLVSVCSAVSHRNQQTWTEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

QY 61 RIHVYTFVQKMPGQLPILGEVDFSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

```
Db 61 RIHVYTFVQKMPGQLPILEGVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
Qy 121 KKTPTVVLKATAGLRLLPEHAKALLFEVKEIFRKSPLVPKGSVINTGODEGIFAWVTV 180
Db 121 KKTPTVVLKATAGLRLLPEHAKALLFEVKEIFRKSPLVPKGSVINTGODEGIFAWVTV 180
Qy 181 NFLTGQLHGRQETVGTGLDGGASTQITFLPQPEKTLQTPRGYLTSEFMFNSTYKLYTH 240
Db 181 NFLTGQLHGRQETVGTGLDGGASTQITFLPQPEKTLQTPRGYLTSEFMFNSTYKLYTH 240
Qy 241 SYLGFGGLKAARLATLGALETGDTGHTFRSACLPRWLEAWIFGGVKYQYGGNQEVEGF 300
Db 241 SYLGFGGLKAARLATLGALETGDTGHTFRSACLPRWLEAWIFGGVKYQYGGNQEVEGF 300
Qy 301 EPCYAEVLVRVGRKLGHOPEEVQSGFYAFSYDDRAVDTDMDIDYKGGILKVEDFERKAR 360
Db 301 EPCYAEVLVRVGRKLGHOPEEVQSGFYAFSYDDRAVDTDMDIDYKGGILKVEDFERKAR 360
Qy 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTKKVNNIETGVALGATFHL 420
Db 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTKKVNNIETGVALGATFHL 420
Qy 421 LOSLGISH 428
Db 421 LOSLGISH 428

RESULT 2
US-09-350-836B-7
; Sequence 7, Application US/09350836B
; Patent No. 6387645
; GENERAL INFORMATION:
; APPLICANT: Mulero, John
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; FILE REFERENCE: 28110/35761
; CURRENT APPLICATION NUMBER: US/09/350,836B
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 7
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-836B-7

Query Match 100.0%; Score 2252; DB 4; Length 428;
Best Local Similarity 100.0%; Pred. No. 7.8e-247;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATSWGTVPFVFLVWVCVSAVSHRNQOTWFEGIFLSSMCPINVSASTLYGINFMDAGSTGT 60
Db 1 MATSWGTVPFVFLVWVCVSAVSHRNQOTWFEGIFLSSMCPINVSASTLYGINFMDAGSTGT 60
Qy 61 RIHVYTFVQKMPGQLPILEGVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
Db 61 RIHVYTFVQKMPGQLPILEGVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
Qy 121 KKTPTVVLKATAGLRLLPEHAKALLFEVKEIFRKSPLVPKGSVINTGODEGIFAWVTV 180
Db 121 KKTPTVVLKATAGLRLLPEHAKALLFEVKEIFRKSPLVPKGSVINTGODEGIFAWVTV 180
Qy 181 NFLTGQLHGRQETVGTGLDGGASTQITFLPQPEKTLQTPRGYLTSEFMFNSTYKLYTH 240
Db 181 NFLTGQLHGRQETVGTGLDGGASTQITFLPQPEKTLQTPRGYLTSEFMFNSTYKLYTH 240
Qy 241 SYLGFGGLKAARLATLGALETGDTGHTFRSACLPRWLEAWIFGGVKYQYGGNQEVEGF 300
```

```
Db 181 NFLTGQLHGRQETVGTGLDGGASTQITFLPQPEKTLQTPRGYLTSEFMFNSTYKLYTH 240
Qy 241 SYLGFGGLKAARLATLGALETGDTGHTFRSACLPRWLEAWIFGGVKYQYGGNQEVEGF 300
Db 241 SYLGFGGLKAARLATLGALETGDTGHTFRSACLPRWLEAWIFGGVKYQYGGNQEVEGF 300
Qy 301 EPCYAEVLVRVGRKLGHOPEEVQSGFYAFSYDDRAVDTDMDIDYKGGILKVEDFERKAR 360
Db 301 EPCYAEVLVRVGRKLGHOPEEVQSGFYAFSYDDRAVDTDMDIDYKGGILKVEDFERKAR 360
Qy 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTKKVNNIETGVALGATFHL 420
Db 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTKKVNNIETGVALGATFHL 420
Qy 421 LOSLGISH 428
Db 421 LOSLGISH 428

RESULT 3
US-09-370-265-7
; Sequence 7, Application US/09370265
; Patent No. 6447771
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; FILE REFERENCE: 28111/35908
; CURRENT APPLICATION NUMBER: US/09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 7
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-370-265-7

Query Match 100.0%; Score 2252; DB 4; Length 428;
Best Local Similarity 100.0%; Pred. No. 7.8e-247;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATSWGTVPFVFLVWVCVSAVSHRNQOTWFEGIFLSSMCPINVSASTLYGINFMDAGSTGT 60
Db 1 MATSWGTVPFVFLVWVCVSAVSHRNQOTWFEGIFLSSMCPINVSASTLYGINFMDAGSTGT 60
Qy 61 RIHVYTFVQKMPGQLPILEGVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
Db 61 RIHVYTFVQKMPGQLPILEGVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
Qy 121 KKTPTVVLKATAGLRLLPEHAKALLFEVKEIFRKSPLVPKGSVINTGODEGIFAWVTV 180
Db 121 KKTPTVVLKATAGLRLLPEHAKALLFEVKEIFRKSPLVPKGSVINTGODEGIFAWVTV 180
Qy 181 NFLTGQLHGRQETVGTGLDGGASTQITFLPQPEKTLQTPRGYLTSEFMFNSTYKLYTH 240
Db 181 NFLTGQLHGRQETVGTGLDGGASTQITFLPQPEKTLQTPRGYLTSEFMFNSTYKLYTH 240
Qy 241 SYLGFGGLKAARLATLGALETGDTGHTFRSACLPRWLEAWIFGGVKYQYGGNQEVEGF 300
```

241 SYLGFLKAARLATIAGALETEGTDGHTFRSACLPRLAEAWIFGGVKYQYGGNQEVEGF 300
241 SYLGFLKAARLATIAGALETEGTDGHTFRSACLPRLAEAWIFGGVKYQYGGNQEVEGF 300
301 EPCYAEVLVRVGRKGLHQPVEVQSGSFYAFSYDDRAVDMDIDYKGGILKVEDFERKAR 360
301 EPCYAEVLVRVGRKGLHQPVEVQSGSFYAFSYDDRAVDMDIDYKGGILKVEDFERKAR 360
301 EPCYAEVLVRVGRKGLHQPVEVQSGSFYAFSYDDRAVDMDIDYKGGILKVEDFERKAR 360
361 EVCNDNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATPHL 420
361 EVCNDNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATPHL 420
421 LQSLGISH 428
421 LQSLGISH 428

RESULT 4
US-09-557-800C-7
; Sequence 7, Application US/09557800C
; Patent No. 6476211
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 28110/36457
; CURRENT APPLICATION NUMBER: US/09/557,800C
; CURRENT FILING DATE: 2000-04-25
; PRIOR FILING DATE: 2000-01-11
; PRIOR FILING DATE: 2000-01-11
; PRIOR FILING DATE: 1999-08-09
; PRIOR FILING DATE: 1999-08-09
; PRIOR FILING DATE: 1999-07-16
; PRIOR FILING DATE: 1999-07-16
; PRIOR FILING DATE: 1999-07-09
; PRIOR FILING DATE: 1999-03-19
; PRIOR FILING DATE: 1999-03-19
; PRIOR FILING DATE: 1998-07-24
; PRIOR FILING DATE: 1998-07-24
; PRIOR FILING DATE: 1999-02-04
; PRIOR FILING DATE: 1999-02-04
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-557-800C-7

Query Match 100.0%; Score 2252; DB 4; Length 428;
Best Local Similarity 100.0%; Pred. No. 7.8e-247;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATSWGTFFMLVWSCVSAVSHRNQOTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
DB 1 MATSWGTFFMLVWSCVSAVSHRNQOTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
QY 61 RIHYVTFVQKMPGQLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
DB 61 RIHYVTFVQKMPGQLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
QY 121 KKTVPVLKATAGLRLPEHKAKALLFEVKEIFRKSPLVPKGSVIMTQDEGIFAWTV 180
DB 121 KKTVPVLKATAGLRLPEHKAKALLFEVKEIFRKSPLVPKGSVIMTQDEGIFAWTV 180
QY 181 NFLTQLGHRQETVGTDLGGASTQITFLPQFETLEQTPRGYLTSEFMFNSTYKLYTH 240
DB 181 NFLTQLGHRQETVGTDLGGASTQITFLPQFETLEQTPRGYLTSEFMFNSTYKLYTH 240

241 SYLGFLKAARLATIAGALETEGTDGHTFRSACLPRLAEAWIFGGVKYQYGGNQEVEGF 300
241 SYLGFLKAARLATIAGALETEGTDGHTFRSACLPRLAEAWIFGGVKYQYGGNQEVEGF 300
301 EPCYAEVLVRVGRKGLHQPVEVQSGSFYAFSYDDRAVDMDIDYKGGILKVEDFERKAR 360
301 EPCYAEVLVRVGRKGLHQPVEVQSGSFYAFSYDDRAVDMDIDYKGGILKVEDFERKAR 360
301 EPCYAEVLVRVGRKGLHQPVEVQSGSFYAFSYDDRAVDMDIDYKGGILKVEDFERKAR 360
361 EVCNDNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATPHL 420
361 EVCNDNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATPHL 420
421 LQSLGISH 428
421 LQSLGISH 428

RESULT 5
US-09-370-625A-7
; Sequence 7, Application US/09370625A
; Patent No. 6600032
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDE
; FILE REFERENCE: 28110/35908
; CURRENT APPLICATION NUMBER: US/09/370,625A
; CURRENT FILING DATE: 1999-08-09
; PRIOR FILING DATE: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR FILING DATE: 1999-07-09
; PRIOR FILING DATE: 1999-07-09
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-370-625A-7

Query Match 100.0%; Score 2252; DB 4; Length 428;
Best Local Similarity 100.0%; Pred. No. 7.8e-247;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATSWGTFFMLVWSCVSAVSHRNQOTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
DB 1 MATSWGTFFMLVWSCVSAVSHRNQOTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
QY 61 RIHYVTFVQKMPGQLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
DB 61 RIHYVTFVQKMPGQLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
QY 121 KKTVPVLKATAGLRLPEHKAKALLFEVKEIFRKSPLVPKGSVIMTQDEGIFAWTV 180
DB 121 KKTVPVLKATAGLRLPEHKAKALLFEVKEIFRKSPLVPKGSVIMTQDEGIFAWTV 180
QY 181 NFLTQLGHRQETVGTDLGGASTQITFLPQFETLEQTPRGYLTSEFMFNSTYKLYTH 240
DB 181 NFLTQLGHRQETVGTDLGGASTQITFLPQFETLEQTPRGYLTSEFMFNSTYKLYTH 240
QY 241 SYLGFLKAARLATIAGALETEGTDGHTFRSACLPRLAEAWIFGGVKYQYGGNQEVEGF 300
DB 241 SYLGFLKAARLATIAGALETEGTDGHTFRSACLPRLAEAWIFGGVKYQYGGNQEVEGF 300
QY 301 EPCYAEVLVRVGRKGLHQPVEVQSGSFYAFSYDDRAVDMDIDYKGGILKVEDFERKAR 360
DB 301 EPCYAEVLVRVGRKGLHQPVEVQSGSFYAFSYDDRAVDMDIDYKGGILKVEDFERKAR 360
QY 361 EVCNDNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATPHL 420
DB 361 EVCNDNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTKKVNNIETGWALGATPHL 420

Db 301 EPCYAEVLVRGKLHQPEEVQSGFYAFSVYVYDRAVDTMDIYKGGILKVEDFERKAR 360
QY 361 EVCNLENTSGSPFLCNDLSYITALLKDGEGFADSTVLQTLTKKNNIETGHALGATPHL 420
Db 361 EVCNLENTSGSPFLCNDLSYITALLKDGEGFADSTVLQTLTKKNNIETGHALGATPHL 420
QY 421 LQSLGISH 428
Db 421 LQSLGISH 428

RESULT 7

US-09-608-285A-5
; Sequence 5, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 28110/36570
; CURRENT APPLICATION NUMBER: US/09/608,285A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/583,231
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-608-285A-5

Query Match 99.2%; Score 2235; DB 4; Length 428;
Best Local Similarity 99.3%; Pred. No. 6.7e-245; Indels 0; Gaps 0;
Matches 425; Conservative 0; Mismatches 3;

QY 1 MATSGTVFFMLVVCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSGTVFFMLVVCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
QY 61 RIHYTVFQKMPGQLPILEGEVDSVKGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
Db 61 RIHYTVFQKMPGQLPILEGEVDSVKGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
QY 121 KKTVPVLKATAGLRLLPEHAKALLFEVKEIFRKSPLVPKGSVSIMTQDEGIPAWTV 180
Db 121 KKTVPVLKATAGLRLLPEHAKALLFEVKEIFRKSPLVPKGSVSIMTQDEGIPAWTV 180
QY 181 NFLTQGLHGRQETVGTLDLGGASTQITFLPQPKTLEQTPRGYLTSPFEMNSTYKLYTH 240
Db 181 NFLTQGLHGRQETVGTLDLGGASTQITFLPQPKTLEQTPRGYLTSPFEMNSTYKLYTH 240
QY 241 SYLGFGLKAARLTIAGLETGDTGHTFRSACLPRWLEAEWIFGGVKYQYGNQGEVGF 300

QY 421 LQSLGISH 428
Db 421 LQSLGISH 428

RESULT 6

US-09-608-285A-3
; Sequence 3, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 28110/36570
; CURRENT APPLICATION NUMBER: US/09/608,285A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/583,231
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-608-285A-3

Query Match 99.2%; Score 2235; DB 4; Length 428;
Best Local Similarity 99.3%; Pred. No. 6.7e-245; Indels 0; Gaps 0;
Matches 425; Conservative 0; Mismatches 3;

QY 1 MATSGTVFFMLVVCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSGTVFFMLVVCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
QY 61 RIHYTVFQKMPGQLPILEGEVDSVKGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
Db 61 RIHYTVFQKMPGQLPILEGEVDSVKGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
QY 121 KKTVPVLKATAGLRLLPEHAKALLFEVKEIFRKSPLVPKGSVSIMTQDEGIPAWTV 180
Db 121 KKTVPVLKATAGLRLLPEHAKALLFEVKEIFRKSPLVPKGSVSIMTQDEGIPAWTV 180
QY 181 NFLTQGLHGRQETVGTLDLGGASTQITFLPQPKTLEQTPRGYLTSPFEMNSTYKLYTH 240
Db 181 NFLTQGLHGRQETVGTLDLGGASTQITFLPQPKTLEQTPRGYLTSPFEMNSTYKLYTH 240
QY 241 SYLGFGLKAARLTIAGLETGDTGHTFRSACLPRWLEAEWIFGGVKYQYGNQGEVGF 300
Db 241 SYLGFGLKAARLTIAGLETGDTGHTFRSACLPRWLEAEWIFGGVKYQYGNQGEVGF 300
QY 301 EPCYAEVLVRGKLHQPEEVQSGFYAFSVYVYDRAVDTMDIYKGGILKVEDFERKAR 360

Db 241 SYLGFLKAARLATIAGLETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQBGEVGF 300
QY 301 EPCYAEVLVRVGRKQHQPVEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 301 EPCYAEVLVRVGRKQHQPVEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
QY 361 EVCMDLENFTSGSPFLCMDLSVITALLKDGFGFADSTVLQTKKNNIETGVALGATFHL 420
Db 361 EVCMDLENFTSGSPFLCMDLSVITALLKDGFGFADSTVLQTKKNNIETGVALGATFHL 420
QY 421 LQSLGISH 428
Db 421 LQSLGISH 428

RESULT 8

US-09-240-639-5
; Sequence 6, Application US/09240639
; Patent No. 6350447
; GENERAL INFORMATION:
; APPLICANT: Chadwick, Brian Paul
; APPLICANT: Frischaut, Anna-Maria
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO CD39-LIKE
; TITLE OF INVENTION: POLYPEPTIDES AND NUCLEIC ACIDS
; FILE REFERENCE: 9598-066
; CURRENT APPLICATION NUMBER: US/09/240,639
; CURRENT FILING DATE: 1998-01-29
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-240-639-5

Query Match 99.2%; Score 2235; DB 4; Length 428;
Best Local Similarity 99.3%; Sred. No. 6.7e-245;
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 MATSMTGTVFFMLVWSCVCSAVSHRNOQTWFGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSMTGTVFFMLVWSCVCSAVSHRNOQTWFGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
QY 61 RIHVTVFVKMPGQQLPILEGVDSVKPGLSAFVDQPKQGAETVQGLLEVAKOSIPRSHW 120
Db 61 RIHVTVFVKMPGQQLPILEGVDSVKPGLSAFVDQPKQGAETVQGLLEVAKOSIPRSHW 120
QY 121 KKTVPVLKATAGLRLLPEHKAKALLFEVKEIFRKSPPFLVPKGSVSIIMTGDGEGIFAWTV 180
Db 121 KKTVPVLKATAGLRLLPEHKAKALLFEVKEIFRKSPPFLVPKGSVSIIMTGDGEGIFAWTV 180
QY 181 NFLTQQLHGRQETVGTLDLGGASTQITFLPQPKETLEQTPRGYLTSPFEMNSTYKLYTH 240
Db 181 NFLTQQLHGRQETVGTLDLGGASTQITFLPQPKETLEQTPRGYLTSPFEMNSTYKLYTH 240
QY 241 SYLGFLKAARLATIAGLETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQBGEVGF 300
Db 241 SYLGFLKAARLATIAGLETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQBGEVGF 300
QY 301 EPCYAEVLVRVGRKQHQPVEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 301 EPCYAEVLVRVGRKQHQPVEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
QY 361 EVCMDLENFTSGSPFLCMDLSVITALLKDGFGFADSTVLQTKKNNIETGVALGATFHL 420
Db 361 EVCMDLENFTSGSPFLCMDLSVITALLKDGFGFADSTVLQTKKNNIETGVALGATFHL 420
QY 421 LQSLGISH 428
Db 421 LQSLGISH 428

RESULT 9

US-09-240-639-9
; Sequence 9, Application US/09240639
; Patent No. 6350447
; GENERAL INFORMATION:
; APPLICANT: Frischaut, Anna-Maria
; APPLICANT: Chadwick, Brian Paul
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO CD39-LIKE
; TITLE OF INVENTION: POLYPEPTIDES AND NUCLEIC ACIDS
; FILE REFERENCE: 9598-066
; CURRENT APPLICATION NUMBER: US/09/240,639
; CURRENT FILING DATE: 1998-01-29
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-240-639-9

Query Match 99.2%; Score 2235; DB 4; Length 428;
Best Local Similarity 99.3%; Pred. No. 6.7e-245;
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 MATSMTGTVFFMLVWSCVCSAVSHRNOQTWFGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSMTGTVFFMLVWSCVCSAVSHRNOQTWFGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
QY 61 RIHVTVFVKMPGQQLPILEGVDSVKPGLSAFVDQPKQGAETVQGLLEVAKOSIPRSHW 120
Db 61 RIHVTVFVKMPGQQLPILEGVDSVKPGLSAFVDQPKQGAETVQGLLEVAKOSIPRSHW 120
QY 121 KKTVPVLKATAGLRLLPEHKAKALLFEVKEIFRKSPPFLVPKGSVSIIMTGDGEGIFAWTV 180
Db 121 KKTVPVLKATAGLRLLPEHKAKALLFEVKEIFRKSPPFLVPKGSVSIIMTGDGEGIFAWTV 180
QY 181 NFLTQQLHGRQETVGTLDLGGASTQITFLPQPKETLEQTPRGYLTSPFEMNSTYKLYTH 240
Db 181 NFLTQQLHGRQETVGTLDLGGASTQITFLPQPKETLEQTPRGYLTSPFEMNSTYKLYTH 240
QY 241 SYLGFLKAARLATIAGLETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQBGEVGF 300
Db 241 SYLGFLKAARLATIAGLETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQBGEVGF 300
QY 301 EPCYAEVLVRVGRKQHQPVEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 301 EPCYAEVLVRVGRKQHQPVEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
QY 361 EVCMDLENFTSGSPFLCMDLSVITALLKDGFGFADSTVLQTKKNNIETGVALGATFHL 420
Db 361 EVCMDLENFTSGSPFLCMDLSVITALLKDGFGFADSTVLQTKKNNIETGVALGATFHL 420
QY 421 LQSLGISH 428
Db 421 LQSLGISH 428

RESULT 10

US-09-350-836B-3
; Sequence 3, Application US/09350836B
; Patent No. 6387645
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 28110/35761
; CURRENT APPLICATION NUMBER: US/09/350,836B
; CURRENT FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16

```
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-836B-3

Query Match          99.2%; Score 2235; DB 4; Length 428;
Best Local Similarity 99.3%; Pred. No. 6.7e-245;
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MATSWGTVFVQKMPGQLPILGEVFDVSKPGLSAFVDPQKQGAETVQGLLEVAKDSIPRSHW 120
DB 1 MATSWGTVFVQKMPGQLPILGEVFDVSKPGLSAFVDPQKQGAETVQGLLEVAKDSIPRSHW 120
QY 61 RIHVYTFVQKMPGQLPILGEVFDVSKPGLSAFVDPQKQGAETVQGLLEVAKDSIPRSHW 120
DB 61 RIHVYTFVQKMPGQLPILGEVFDVSKPGLSAFVDPQKQGAETVQGLLEVAKDSIPRSHW 120
QY 121 KKTTPVVKATAGLRLPEHKAKALLPEVKEIFRKSPLVPKGSVSIMTQDEGIFAWVTV 180
DB 121 KKTTPVVKATAGLRLPEHKAKALLPEVKEIFRKSPLVPKGSVSIMTQDEGIFAWVTV 180
QY 181 NFLTQGLHGRQETVGTLDLGGASTQITFLPQPEKTLQTPRGYLTSPFEMNSTYKLYTH 240
DB 181 NFLTQGLHGRQETVGTLDLGGASTQITFLPQPEKTLQTPRGYLTSPFEMNSTYKLYTH 240
QY 241 SYLPGFLKAARLATIAGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300
DB 241 SYLPGFLKAARLATIAGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300
QY 301 EPCYAEVLVRVGRKLGHPQBEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDPFRKAR 360
DB 301 EPCYAEVLVRVGRKLGHPQBEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDPFRKAR 360
QY 361 EVCNDNLENTSGSPFLCNDLSYITALLKDGFGFADSTVLQTKKVNNIETGWLGAFTFHL 420
DB 361 EVCNDNLENTSGSPFLCNDLSYITALLKDGFGFADSTVLQTKKVNNIETGWLGAFTFHL 420
QY 421 LQSLGISH 428
DB 421 LQSLGISH 428

RESULT 12
US-09-370-265-3
; Sequence 3, Application US/09370265
; Patent No. 6447771
; GENERAL INFORMATION:
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; FILE REFERENCE: 28111/35908
; CURRENT APPLICATION NUMBER: US/09/370,265
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: PCT/US99/16180
; EARLIER FILING DATE: 1999-07-16
; EARLIER APPLICATION NUMBER: 09/350,836
; EARLIER FILING DATE: 1999-07-09
; EARLIER APPLICATION NUMBER: 09/273,447
; EARLIER FILING DATE: 1999-03-19
; EARLIER APPLICATION NUMBER: 09/244,444
; EARLIER FILING DATE: 1999-02-04
; EARLIER APPLICATION NUMBER: 09/122,449
; EARLIER FILING DATE: 1998-07-24
; EARLIER APPLICATION NUMBER: 09/118,205
; EARLIER FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-370-265-3

; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-836B-5

Query Match          99.2%; Score 2235; DB 4; Length 428;
Best Local Similarity 99.3%; Pred. No. 6.7e-245;
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MATSWGTVFVQKMPGQLPILGEVFDVSKPGLSAFVDPQKQGAETVQGLLEVAKDSIPRSHW 120
DB 1 MATSWGTVFVQKMPGQLPILGEVFDVSKPGLSAFVDPQKQGAETVQGLLEVAKDSIPRSHW 120
QY 61 RIHVYTFVQKMPGQLPILGEVFDVSKPGLSAFVDPQKQGAETVQGLLEVAKDSIPRSHW 120
DB 61 RIHVYTFVQKMPGQLPILGEVFDVSKPGLSAFVDPQKQGAETVQGLLEVAKDSIPRSHW 120
QY 121 KKTTPVVKATAGLRLPEHKAKALLPEVKEIFRKSPLVPKGSVSIMTQDEGIFAWVTV 180
DB 121 KKTTPVVKATAGLRLPEHKAKALLPEVKEIFRKSPLVPKGSVSIMTQDEGIFAWVTV 180
QY 181 NFLTQGLHGRQETVGTLDLGGASTQITFLPQPEKTLQTPRGYLTSPFEMNSTYKLYTH 240
DB 181 NFLTQGLHGRQETVGTLDLGGASTQITFLPQPEKTLQTPRGYLTSPFEMNSTYKLYTH 240
QY 241 SYLPGFLKAARLATIAGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300
DB 241 SYLPGFLKAARLATIAGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300
QY 301 EPCYAEVLVRVGRKLGHPQBEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDPFRKAR 360
DB 301 EPCYAEVLVRVGRKLGHPQBEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDPFRKAR 360
QY 361 EVCNDNLENTSGSPFLCNDLSYITALLKDGFGFADSTVLQTKKVNNIETGWLGAFTFHL 420
DB 361 EVCNDNLENTSGSPFLCNDLSYITALLKDGFGFADSTVLQTKKVNNIETGWLGAFTFHL 420
QY 421 LQSLGISH 428
DB 421 LQSLGISH 428

RESULT 11
US-09-350-836B-5
; Sequence 5, Application US/09350836B
; Patent No. 6387645
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; FILE REFERENCE: 28110/35761
; CURRENT APPLICATION NUMBER: US/09/350,836B
; CURRENT FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 428
```

Query Match 99.2%; Score 2235; DB 4; Length 428;
 Best Local Similarity 99.3%; Pred. No. 6.7e-245; Indels 0; Gaps 0;
 Matches 425; Conservative 0; Mismatches 3;

QY 1 MATSWGTVPFVWVSVCSAVSHRNQQTWFGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
 DB 1 MATSWGTVPFVWVSVCSAVSHRNQQTWFGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

QY 61 RIHVYTFVQKMPGOLPILLEGVDSVKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120
 DB 61 RIHVYTFVQKMPGOLPILLEGVDSVKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120

QY 121 KCTPVVLKATAGRLRLPEHAKALLFEVKEIPRKSPLVPKGSVIMTGDGEGIFAWTV 180
 DB 121 KCTPVVLKATAGRLRLPEHAKALLFEVKEIPRKSPLVPKGSVIMTGDGEGIFAWTV 180

QY 181 NFLTQGLHGRHOBTVGTLDLGGASTQITFLPQFEXTLEOTPRGYLTSFEMFNSTYKLYTH 240
 DB 181 NFLTQGLHGRHOBTVGTLDLGGASTQITFLPQFEXTLEOTPRGYLTSFEMFNSTYKLYTH 240

QY 241 SYLGFGLKAARLALATLGALETGTDGHTFRSACLPRMLEAEWIFGGVKYQYGGNQEVEVF 300
 DB 241 SYLGFGLKAARLALATLGALETGTDGHTFRSACLPRMLEAEWIFGGVKYQYGGNQEVEVF 300

QY 301 EPCYAEVLVRVGRKGLHQPVEVQSGFYAFSYYYDRAVDTMDIYKGGILKVEDFERKAR 360
 DB 301 EPCYAEVLVRVGRKGLHQPVEVQSGFYAFSYYYDRAVDTMDIYKGGILKVEDFERKAR 360

QY 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQTKKVNNIETGVALGATFHL 420
 DB 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQTKKVNNIETGVALGATFHL 420

QY 421 LQSLGISH 428
 DB 421 LQSLGISH 428

RESULT 13
 US-09-370-265-5
 ; Sequence 5, Application US/09370265
 ; Patent No. 6447771
 ; GENERAL INFORMATION:
 ; APPLICANT: Ford, John
 ; APPLICANT: Mulero, Julio
 ; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
 ; FILE REFERENCE: 28111/35908
 ; CURRENT APPLICATION NUMBER: US/09/370,265
 ; EARLIER FILING DATE: 1999-08-09
 ; EARLIER FILING DATE: 1999-07-16
 ; EARLIER FILING DATE: 1999-07-16
 ; EARLIER FILING DATE: 1999-07-09
 ; EARLIER FILING DATE: 1999-07-09
 ; EARLIER FILING DATE: 1999-03-19
 ; EARLIER FILING DATE: 1999-03-19
 ; EARLIER FILING DATE: 1999-02-04
 ; EARLIER FILING DATE: 1998-07-24
 ; EARLIER FILING DATE: 1998-07-24
 ; EARLIER FILING DATE: 1998-07-16
 ; NUMBER OF SEQ ID NOS: 37
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 428
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-370-265-5

Query Match 99.2%; Score 2235; DB 4; Length 428;
 Best Local Similarity 99.3%; Pred. No. 6.7e-245; Indels 0; Gaps 0;
 Matches 425; Conservative 0; Mismatches 3;

QY 1 MATSWGTVPFVWVSVCSAVSHRNQQTWFGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
 DB 1 MATSWGTVPFVWVSVCSAVSHRNQQTWFGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

QY 61 RIHVYTFVQKMPGOLPILLEGVDSVKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120
 DB 61 RIHVYTFVQKMPGOLPILLEGVDSVKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120

QY 121 KCTPVVLKATAGRLRLPEHAKALLFEVKEIPRKSPLVPKGSVIMTGDGEGIFAWTV 180
 DB 121 KCTPVVLKATAGRLRLPEHAKALLFEVKEIPRKSPLVPKGSVIMTGDGEGIFAWTV 180

QY 181 NFLTQGLHGRHOBTVGTLDLGGASTQITFLPQFEXTLEOTPRGYLTSFEMFNSTYKLYTH 240
 DB 181 NFLTQGLHGRHOBTVGTLDLGGASTQITFLPQFEXTLEOTPRGYLTSFEMFNSTYKLYTH 240

QY 241 SYLGFGLKAARLALATLGALETGTDGHTFRSACLPRMLEAEWIFGGVKYQYGGNQEVEVF 300
 DB 241 SYLGFGLKAARLALATLGALETGTDGHTFRSACLPRMLEAEWIFGGVKYQYGGNQEVEVF 300

QY 301 EPCYAEVLVRVGRKGLHQPVEVQSGFYAFSYYYDRAVDTMDIYKGGILKVEDFERKAR 360
 DB 301 EPCYAEVLVRVGRKGLHQPVEVQSGFYAFSYYYDRAVDTMDIYKGGILKVEDFERKAR 360

QY 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQTKKVNNIETGVALGATFHL 420
 DB 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQTKKVNNIETGVALGATFHL 420

QY 421 LQSLGISH 428
 DB 421 LQSLGISH 428

RESULT 14
 US-09-557-800C-3
 ; Sequence 3, Application US/09557800C
 ; Patent No. 6476211
 ; GENERAL INFORMATION:
 ; APPLICANT: Ford, John
 ; APPLICANT: Mulero, Julio
 ; APPLICANT: Yeung, George
 ; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
 ; FILE REFERENCE: 28110/36457
 ; CURRENT APPLICATION NUMBER: US/09/557,800C
 ; CURRENT FILING DATE: 2000-04-25
 ; PRIOR FILING DATE: 2000-01-11
 ; PRIOR FILING DATE: 1999-08-09
 ; PRIOR FILING DATE: 1999-08-09
 ; PRIOR FILING DATE: 1999-07-16
 ; PRIOR FILING DATE: 1999-07-16
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR FILING DATE: 1999-07-09
 ; PRIOR FILING DATE: 1999-03-19
 ; PRIOR FILING DATE: 1999-03-19
 ; PRIOR FILING DATE: 1998-07-24
 ; PRIOR FILING DATE: 1998-07-24
 ; PRIOR FILING DATE: 1999-02-04
 ; PRIOR FILING DATE: 1999-02-04
 ; PRIOR FILING DATE: 1998-07-16
 ; NUMBER OF SEQ ID NOS: 56
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 428
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-557-800C-3

Query Match 99.2%; Score 2235; DB 4; Length 428;
 Best Local Similarity 99.3%; Pred. No. 6.7e-245; Indels 0; Gaps 0;
 Matches 425; Conservative 0; Mismatches 3;

```

QY 1 MATSGTVPFVPMVSVCSAVSHRNQOTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
DB 1 MATSGTVPFVPMVSVCSAVSHRNQOTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
QY 61 RIHVYTFVQKMPGQOLPILEGVEFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
DB 61 RIHVYTFVQKMPGQOLPILEGVEFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
QY 121 KKTPTVWLKATAGRLRLPEHKAALLFEVKEIFRKSPLFVPGKSVSINTGQDEGIFAWVTV 180
DB 121 KKTPTVWLKATAGRLRLPEHKAALLFEVKEIFRKSPLFVPGKSVSINTGQDEGIFAWVTV 180
QY 181 NFLTQGLHGRQETVGTLDLGGASTQITFLPQFEKTLQOTPRGYLTSFEMFNSTYKLYTH 240
DB 181 NFLTQGLHGRQETVGTLDLGGASTQITFLPQFEKTLQOTPRGYLTSFEMFNSTYKLYTH 240
QY 241 SYLGFGGLKAARLATIAGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300
DB 241 SYLGFGGLKAARLATIAGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300
QY 301 EPCYAEVLVRVRGKLHQBEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
DB 301 EPCYAEVLVRVRGKLHQBEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
QY 361 EVCNDLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQLTKKVNNIETGALGATPHL 420
DB 361 EVCNDLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQLTKKVNNIETGALGATPHL 420
QY 421 LQSLGISH 428
DB 421 LQSLGISH 428

```

```

RESULT 15
US-09-557-800C-5
; Sequence S, Application US/09557800C
; Patent No. 6476211
; GENERAL INFORMATION:
; APPLICANT: Mulero, John
; APPLICANT: Yeung, George
; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 28110/36457
; CURRENT APPLICATION NUMBER: US/09/557,800C
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/122449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-557-800C-5

```

Query Match 99.2%; Score 2235; DB 4; Length 428;
 Best Local Similarity 99.3%; Pred. No. 6.7e-245;

```

Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 MATSGTVPFVPMVSVCSAVSHRNQOTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
DB 1 MATSGTVPFVPMVSVCSAVSHRNQOTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
QY 61 RIHVYTFVQKMPGQOLPILEGVEFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
DB 61 RIHVYTFVQKMPGQOLPILEGVEFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
QY 121 KKTPTVWLKATAGRLRLPEHKAALLFEVKEIFRKSPLFVPGKSVSINTGQDEGIFAWVTV 180
DB 121 KKTPTVWLKATAGRLRLPEHKAALLFEVKEIFRKSPLFVPGKSVSINTGQDEGIFAWVTV 180
QY 181 NFLTQGLHGRQETVGTLDLGGASTQITFLPQFEKTLQOTPRGYLTSFEMFNSTYKLYTH 240
DB 181 NFLTQGLHGRQETVGTLDLGGASTQITFLPQFEKTLQOTPRGYLTSFEMFNSTYKLYTH 240
QY 241 SYLGFGGLKAARLATIAGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300
DB 241 SYLGFGGLKAARLATIAGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300
QY 301 EPCYAEVLVRVRGKLHQBEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
DB 301 EPCYAEVLVRVRGKLHQBEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
QY 361 EVCNDLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQLTKKVNNIETGALGATPHL 420
DB 361 EVCNDLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQLTKKVNNIETGALGATPHL 420
QY 421 LQSLGISH 428
DB 421 LQSLGISH 428

```

Search completed: July 1, 2004, 13:50:11
 Job time : 17.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2004, 13:42:31 ; Search time 10 Seconds
(without alignments)
2228.603 Million cell updates/sec

Title: US-10-091-085-7

Perfect score: 2252

Sequence: 1 MATSMGTVPFMLVYVCVCSA.....ETGVALGATFHLQLSGISH 428

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2235	99.2	428	1	ENP5 HUMAN
2	1979.5	87.9	427	1	ENP5 MOUSE
3	1918.5	85.2	469	1	ENP5 MESAU
4	996	44.2	484	1	ENP6 HUMAN
5	983	43.7	455	1	ENP6 RAT
6	488.5	21.7	518	1	GDAL YEAST
7	484	21.5	455	1	NTPA_PEA
8	441.5	19.6	454	1	APY_SOUTU
9	394	17.5	493	1	ENP1 CHICK
10	387	17.2	552	1	YBU4 CAEEL
11	385	17.1	510	1	ENP1 MOUSE
12	382.5	17.0	485	1	YV4E CAEEL
13	375	16.7	513	1	ENP1 BOVIN
14	371	16.5	630	1	YNDT YEAST
15	367	16.3	510	1	ENP1 PIG
16	366.5	16.3	613	1	ENP4 MOUSE
17	366	16.3	510	1	ENP1 HUMAN
18	363.5	16.1	511	1	ENP1 RAT
19	363	16.1	616	1	ENP4 HUMAN
20	357.5	15.9	529	1	ENP3 HUMAN
21	351.5	15.6	495	1	ENP2 MOUSE
22	321.5	14.3	495	1	ENP2 RAT
23	314.5	14.0	495	1	ENP2 CHICK
24	310.5	13.8	494	1	ENP2 BOVIN
25	150	6.7	628	1	NTP1 TOXGO
26	148	6.6	628	1	NTP2 TOXGO
27	141	6.3	592	1	NTP4 TOXGO
28	105.5	4.7	533	1	VGL SPV4
29	104.5	4.6	535	1	DEXB STRPN
30	100.5	4.5	634	1	SELB MOOTH
31	100.5	4.5	1151	1	Y245 TREPA
32	99	4.4	711	1	DPF3 YEAST
33	98.5	4.4	4427	1	PKSL_BACSU

ALIGNMENTS

```
RESULT 1
ENP5_HUMAN
ID ENP5_HUMAN STANDARD; PRT; 428 AA.
AC O75356;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 5 precursor
DE (EC 3.6.1.6) (NTPDase5) (Nucleoside diphosphatase) (CD39 antigen-like
DE 4) (ER-UDPase).
GN NTPD5 OR CD39L4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leukemia;
RX MEDLINE=98341119; PubMed=9676430;
RA Chadwick B.P., Frischauf A.-M.;
RT "The CD39-like gene family: identification of three new human members
RT (CD39L2, CD39L3, and CD39L4), their murine homologues, and a member of
RT the gene family from Drosophila melanogaster.";
RL Genomics 50:357-367(1998)
CC -!- FUNCTION: Likely to promote reglycosylation reactions involved in
CC glycoproteins folding and quality control in the endoplasmic
CC reticulum. Hydrolyzes UDP, GDP AND IDP but not any other
CC nucleoside di-, mono- or triphosphates, nor thiamine pyrophosphate
CC (By similarity).
CC -!- CATALYTIC ACTIVITY: A nucleoside diphosphate + H(2)O = a
CC nucleotide + phosphate.
CC -!- COFACTOR: Requires calcium and magnesium.
CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen (By similarity).
CC -!- TISSUE SPECIFICITY: Expressed in adult liver, kidney, prostate,
CC testis and colon. Much weaker expression in other tissues.
CC -!- SIMILARITY: Belongs to the GDAL / CD39 NTPase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF039918; AAC39885.1; -.
CC Genew; HGNC:3367; ENTPD5.
CC MIM; 603162; -.
CC InterPro; IPR000407; GDAL_CD39_NTPase.
CC Pfam; PF01150; GDAL_CD39; 1.
CC PROSITE; PS01238; GDAL_CD39_NTPASE; FALSE NEG.
CC Hydroxylase; Transmembrane; Glycoprotein; Calcium; Magnesium;
CC Endoplasmic reticulum; Signal.
CC SIGNAL 1 20 POTENTIAL.
CC CHAIN 21 428 ECTONUCLEOSIDE TRIPHOSPHATE
CC DIPIHOSPHOHYDROLASE 5.
```

```
34 96.5 4.3 287 1 RM07 SCHPO
35 96 4.3 477 1 URT1 DESRO
36 95.5 4.2 301 1 PSD CHLTR
37 92 4.1 290 1 HEM6 CAUCR
38 92 4.1 424 1 Y115 ARCFU
39 92 4.1 1220 1 PTC1 ERASE
40 92 4.1 1549 1 Y1R3 YEAST
41 92 4.1 1549 1 YOW2 YEAST
42 91.5 4.1 796 1 TLR6 HUMAN
43 91 4.0 417 1 GLYA BORBU
44 91 4.0 946 1 ITH2 MOUSE
45 90.5 4.0 279 1 HEM6 RICPR
014337 schizosacch
P98119 desmodus ro
O84705 chlamydia t
Q9aat8 caulobacter
P56969 archaeoglob
Q98864 brachydanio
P40438 saccharomyc
P40890 saccharomyc
Q9Y2C9 homo sapien
O51547 borrelia bu
Q61703 mus musculu
Q92C86 rickettsia
```

```
FT CARBOHYD 42 42 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 232 232 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 368 368 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 428 AA; 47517 MW; 830437A155D84DD CRC64;

Query Match 99.2%; Score 2235; DB 1; Length 428;
Best Local Similarity 99.3%; Pred. No. 1.2e-170;
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MATSGTGVFFMLVSVCSAVSHRNQQTWTFEGIFLSSMPCPNVNSASTLYGIMFDAGSTGT 60
D5 1 MATSGTGVFFMLVSVCSAVSHRNQQTWTFEGIFLSSMPCPNVNSASTLYGIMFDAGSTGT 60
QY 61 RIHYVTFVQKPGQPLILEGVFDSVKPGLSAFVDQPKQAGTQVQGLLEVAKDSIPRSHW 120
D5 61 RIHYVTFVQKPGQPLILEGVFDSVKPGLSAFVDQPKQAGTQVQGLLEVAKDSIPRSHW 120
QY 121 KKTPTVLKATAGLRLPEHAKALLFEVKELFRKSPFLVPKGSVIMTQDEGFEAWTV 180
D5 121 KKTPTVLKATAGLRLPEHAKALLFEVKELFRKSPFLVPKGSVIMTQDEGFEAWTV 180
QY 181 NFLGQLHGHQETVGTLDLGASTQITFLPQFKTLEQTPRGYLTSPFMENSTYKLYTH 240
D5 181 NFLGQLHGHQETVGTLDLGASTQITFLPQFKTLEQTPRGYLTSPFMENSTYKLYTH 240
QY 241 SYLGFGLKAARLATIGALETGTGHTFRSACLPRWLEAEWIFGVKYQVGNQEGVGF 300
D5 241 SYLGFGLKAARLATIGALETGTGHTFRSACLPRWLEAEWIFGVKYQVGNQEGVGF 300
QY 301 EPCYAEVLVRVGRKLHQPEEYQVGSFYAFSTYYDRAVDTMDIDYKGGILKVEDFERKAR 360
D5 301 EPCYAEVLVRVGRKLHQPEEYQVGSFYAFSTYYDRAVDTMDIDYKGGILKVEDFERKAR 360
QY 361 EVCNLENFTSGSPFLCNDLSTYITALLKDXGFGADSTVLQTKKVNITGWLGAFTPL 420
D5 361 EVCNLENFTSGSPFLCNDLSTYITALLKDXGFGADSTVLQTKKVNITGWLGAFTPL 420
QY 421 LQSLGISH 428
D5 421 LQSLGISH 428

RESULT 2
ID ENPS MOUSE STANDARD; PRT; 427 AA.
AC Q9WU29; O70214;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 5 precursor
DE (EC 3.6.1.6) (NTPDase5) (Nucleoside diphosphatase) (CD39 antigen-like
DE 4) (ER-UDPase).
GN ENTPD5 OR CD39L4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=98119025; PubMed=9457681;
RX Chadwick B.P., Williamson J., Sheer D., Frischauf A.-M.;
RT "CDNA cloning and chromosomal mapping of a mouse gene with homology to
RT NTPases.";
RL Mamm. Genome 9:162-164 (1998).
[2]
RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND CHARACTERIZATION.
RC TISSUE=Liver;
EX MEDLINE=99298181; PubMed=10369669;
RA Trombetta B.S., Helenius A.;
RT "Glycoprotein reglucosylation and nucleotide sugar utilization in the
RT endoplasmic reticulum.";
```

```
RL EMBO J. 18:3282-3292 (1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Komio H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Kamanaka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli P., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsch G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaeerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,
RA Wuyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohseki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
[4]
RN SEQUENCE FROM N.A.
RP STRAIN=FVB/N. TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.J.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Kulyk S.W.,
RA Villalón D.K., Muzny D.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc Natl Acad Sci U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: Likely to promote reglucosylation reactions involved in
CC glycoproteins folding and quality control in the endoplasmic
CC reticulum. Hydrolyzes UDP, GDP AND IDP but not any other
CC nucleoside di-, mono- or triphosphates, nor thiamine
CC pyrophosphate.
CC -!- CATALYTIC ACTIVITY: A nucleoside diphosphate + H(2)O = a
CC nucleotide + phosphate.
CC -!- COFACTOR: Requires calcium and magnesium.
CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC -!- TISSUE SPECIFICITY: Ubiquitous.
CC -!- PTM: Glycosylated with high mannose N-linked glycans.
CC -!- MISCELLANEOUS: Optimal pH is neutral.
CC -!- SIMILARITY: Belongs to the GDA1 / CD39 NTPase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-aib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: AF006482; AAC05181.1; -.
CC EMBL: AJ238636; CAB45533.1; -.
DR
```


RT (aprase) from potato tubers (Solanum tuberosum).";
RL Biochem. Biophys. Res. Commun. 218:916-923(1996).
RN [2]
RP SEQUENCE OF 42-54; 68-95 AND 236-253.
RC STRAIN: cv. Désirée;
RX MEDLINE=9635615; PubMed=8703025;
RA Vasconcelos E.G., Ferreira S.T., de Carvalho T.M.U., de Souza W.,
Kettlun A.M., Mancilla M., Valenzuela M.A., Verjovski-Almeida S.;
RT "Partial purification and immunohistochemical localization of ATP
diphosphohydrolase from Schistosoma mansoni. Immunological cross-
reactivities with potato apyrase and Toxoplasma gondii nucleoside
triphosphate hydrolase.";
RL J. Biol. Chem. 271:22139-22145(1996).
CC -!- FUNCTION: Catalyzes the hydrolysis of phosphoanhydride bonds of
nucleoside tri- and di-phosphates.
CC -!- CATALYTIC ACTIVITY: ATP + 2 H₂O = AMP + 2 phosphate.
CC -!- COFACTOR: Calcium.
CC -!- SUBCELLULAR LOCATION: Membrane-associated (Probable).
CC -!- PTM: The N-terminus is blocked.
CC -!- SIMILARITY: Belongs to the GDA1 / CD39 NTPase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U58597; AB02720.1; -;
DR PIR; JC4616;
DR InterPro; IPR000407; GDA1_CD39_NTPase.
DR Pfam; PF01150; GDA1_CD39; 1; NTPase; 1.
DR PROSITE; PS01238; GDA1_CD39_NTPase; 1.
KW Hydrolase; Transmembrane; Calcium; Signal.
FT SIGNAL 1 30 POTENTIAL.
FT CHAIN 31 454 APYRASE.
FT TRANSMEM 426 446 POTENTIAL.
FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 454 AA; 50041 MW; 909EF431DAF52P CRC64;

Query Match 19.6%; Score 441.5; DB 1; Length 454;
Best Local Similarity 28.6%; Pred. No. 1.2e-27;
Matches 130; Conservative 78; Mismatches 178; Indels 69; Gaps 16;

Qy 25 NQQTFFGIGFLS-----SMCPINVS-----STLIGMFDAGSTGTRHYV 65
Db 3 NQNGHPIFILAFLPLSLSKKNVAQIPLRHLHSHSEHYAVIFDAGSTGRVHF 62

Qy 66 TFVQKMPQLPILSG-EVFDVSKPGLGAFVDPQKGAETVGLLEVAKDSIPRSHWKTIP 124
Db 63 RFDEKL-GLLPIGNNIEYFWATEPGLSYAEDPKAANSLEPLDGAEGVVPQLQSETP 121

Qy 125 VLKATAGRLRLPEHKAVALFTVKEIFR-KSPFLVPKGSVSIWQDQEGIFAWTVNFI 183
Db 122 LELGATAGRLMKGADEAKILQAVNLVQNGQSTFHSDQWVITLDGTQEGSYNMAATNYL 181

Qy 184 TGLHGHROEVNGLTLDGGASTQTTEL---POPEKTLLE-OTPRGYLTSFEMFNSTYKLYT 239
Db 182 LGNLGDKYKSTATIDLOGGSQVAYALSIEQFAKAPQNEDEGFYVQOKHLMKSDNYLYV 241

Qy 240 HSYLGFGLKAARLATLGALETGTDTGHTFRSACLPRWLEAEWIFGGYKYQVGNQGEVGV 299
Db 242 HSYLYNQLAGRAIFKASRNES-----NPCALEGCDGYSYGGVYKAPKPGS-- 292

Qy 300 FEPYCAEVLVRVG--KLHQPEEVQVGSF-----YAFSYVYDRAVDMDI 342
Db 293 ----SWKFCRRLTRHALKINAKNIECTFNGVMNGGDDGQKRNHASSFFYDIAQGVIV 349

Qy 343 DYE-KGGILKVEDPERKAREVCD-NLENFTS-----GSPFLCMLDSLITALLKDGFG 392
Db 350 DTRKPSALAKPIQVNAAKVACQTNVADIKSIFPKTQDRNIPYLCMDLIVEYLLVDGFG 409

Qy 393 P---ADSTVLQTKKVVN-IETGVALCATPHLQS 423
Db 410 LNPHEKITVHDVQYKNVGVGAAMPGLGCAIDLVSS 444

RESULT 9
ID ENPI_CHICK STANDARD; PRT; 493 AA.
AC 093295;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2003 (Rel. 42, Last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 1 (EC 3.6.1.5)
DE (NTPDase1) [Ecto-ATP diphosphohydrolase (ATPDase) (lymphoid cell
DE activation antigen) (Ecto-apyrase) (CD39 antigen)].
GN ENTPD1 OR CD39.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-21 AND 150-156.
RC TISSUE=Oviduct;
RX MEDLINE=98298108; PubMed=9632655;
RA Nagy A.K., Knowles A.F., Nagami G.T.;
RT "Molecular cloning of the chicken oviduct ecto-APP-
diphosphohydrolase.";
RL J. Biol. Chem. 273:16043-16049(1998).
RN [2]
RP SEQUENCE OF 1-17.
RC TISSUE=Stomach;
RX MEDLINE=97442428; PubMed=9295305;
RA Lewis-Carl S., Kirley T.L.;
RT "Immunolocalization of the ecto-ATPase and ecto-apyrase in chicken
RT gizzard and stomach. Purification and N-terminal sequence of the
RT stomach ecto-apyrase.";
RL J. Biol. Chem. 272:23645-23652(1997).
CC -!- FUNCTION: In the nervous system, could hydrolyze ATP and other
CC nucleotides to regulate purinergic neurotransmission. Could also
CC be implicated in the prevention of platelet aggregation.
CC Hydrolyzes ATP and ADP equally well (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + 2 H₂O = AMP + 2 phosphate.
CC -!- COFACTOR: Requires calcium and magnesium (By similarity).
CC -!- SUBUNIT: Homodimer; disulfide-linked (Probable).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: Belongs to the GDA1 / CD39 NTPase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF041355; AAC26491.1; -;
DR InterPro; IPR000407; GDA1_CD39_NTPase.
DR Pfam; PF01150; GDA1_CD39; 1.
DR PROSITE; PS01238; GDA1_CD39_NTPase; 1.
KW Hydrolase; Transmembrane; Antigen; Glycoprotein; Calcium; Magnesium.
FT DOMAIN 1 7 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 8 28 POTENTIAL.
FT DOMAIN 29 463 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 464 486 POTENTIAL.
FT DOMAIN 487 493 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 65 65 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 79 79 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 223 223 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 267 267 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 324 324 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 361 361 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 372 372 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 16 16 C -> W (IN REP. 21).
FT CONFLICT 21 21 I -> G (IN REP. 1; AA SEQUENCE).
SQ SEQUENCE 493 AA; 54034 MW; F14FF4C3AA2P3603 CRC64;

Query Match 17.5%; Score 394; DB 1; Length 493;
Best Local Similarity 27.8%; Pred. No. 8.5e-24;
Matches 133; Conservative 65; Mismatches 169; Indels 112; Gaps 18;

QY 11 MLVSVCSVSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGTRIHYVTF-VQ 69
Db 11 LTTATCFSTIALILSAVDVXDFV-----PPGKYGLVFDAGSTHTALYVQPAD 62
QY 70 KPGQLPILGEVFSVKGSLSAFVDQPKQGAFTVQGLLEVAKDSIPRSHWKTTPVVLKA 129
Db 63 KENGTGIVSQVESCTVNGSIGSSYADDPAGAGASLKPCLDKAMAVIPVEQQWQPTYLGA 122
QY 130 TAGRLLEPEH---KAKALLEVEKEIFRSPFLVPKGSVSIIMTSGDEGIFAWTVNFLTQ 186
Db 123 TAGKRLLEQNSTKABQVFAVSKAIRFP--VDFRGAQILTCNBSFGHITVNYLLET 180
QY 187 L-----HGRQETVGTLDGGASTQTFLPQPEKTLQTPRGYLSFEMFNSTYK 236
Db 181 LIKFSACKWEHPQNTVEVLGALDGGASTQTFOGV--TIEDKNTSVL--FRLYGTNYS 236
QY 237 LYTHSYLGFGLKAARLATLGALETGDTGTFRSACLPRWLEAEWIFGVKYQYGGQOE- 295
Db 237 LYTHSYLGYGOIQASKRLMAALHQDGSYVQNIHPCYPK-----GVRRI 280
QY 296 ---GEVGPEPCY-----AEVLRVVR-----GKLHQ 317
Db 281 ITIAIYDSCPVEPTSMLSPAQILVTGTGNPAACPTAILKLFNLTCGANTCGFDGVYQ 340
QY 318 PEEVORGSAFYAFS-YYDRA-----VDTMDY-EKGGILKVEDFERKAREVC 363
Db 341 PP--VRGQFPAPAGFYFTFSFLNTGQOSLSHVNATVDFCNKWSSELVETFPQNK----- 394
QY 364 DNLENFTSGSPFLCNDLSYITALLKDGFGFADST--VLQLAKVNNIETGHWALGATPHL 420
Db 395 EHLRTY-----CVGLYILTLLVDGYKFEHTWSNIHFQSKAGNADIGTGLFMLNL 446

RESULT 10
YBU4 CAEEL STANDARD; PRT; 552 AA.
ID YBU4 CAEEL
AC Q21615;
DT 01-NOV-1997 (Rel. 35, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein R07E4.4 in chromosome X.
GN R07E4.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Chordata; Rhabditiida; Rhabditioides;
OC Rhabditiidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Miller N.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RA Waterston R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the GDA1 / CD39 NTPase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U39652; AAA80403.2; -;
DR WormPep; R07E4.4; CE28748.
DR InterPro; IPR000407; GDA1 CD39_NTPase.
DR Pfam; PF01150; GDA1 CD39_1.
DR PROSITE; PS01238; GDA1 CD39_NTPase; 1.
KW Hypothetical protein; Transmembrane; Hydrolase.
FT TRANSKEM 7 27 POTENTIAL.
FT TRANSKEM 490 510 POTENTIAL.
SQ SEQUENCE 552 AA; B5DC32C858AE4D94 CRC64;

Query Match 17.2%; Score 387; DB 1; Length 552;
Best Local Similarity 28.6%; Pred. No. 3.6e-23;
Matches 126; Conservative 70; Mismatches 159; Indels 86; Gaps 20;

QY 49 YGIMFDAGSTGTRIHYVTFVQKMGQGLPILGEVFS-----VKPGLSAFVDQPKQAE 102
Db 44 YGVICDAGSTGTRALFYNNWISTDSBLIQTIEPVIYDNKPFMKKISFGLSTFGTKPAQAAE 103
QY 103 TVQGLLEVAKDSIPRSHWKTTPVVLKATAGRLLEPEHKAALPRVKEIFRKSPPFL---- 158
Db 104 YLRPELEAERHPISEKRPYTPVPIFATAGRLLPDEQEAUL---KLENKLPKITSMQ 160
QY 159 VPKGSVSIIMTSGDEGIFAWTVNFLTQGLH-----CH-ROETVGTLDGGAST 205
Db 161 VLKHEIRITEGKEGIVSMIAVYALGKFNKTATLDPGTSFAHARQKTVGMIDMGASA 220
QY 206 QITF-LPQPEK-----TLEQTPRGYLSFEMFNSTYKLYTHSYLGF---GLKAARLATLG 256
Db 221 QIAFELPDTDFSGSINVENINLGCREDDSLFK--YKLFTVTLGYGNEGIRKYEHMLLS 278
QY 257 ALATEGTGDTGTFRSACLPRWLEAEWIFGVKYQYQGN--QEG-----E 297
Db 279 KLADQ--NGTVIQDDCMPLNLHKT-----VTLENGENFVRRTGNTCSNEVKCLINPE 331
QY 298 VGEPCVAEVLVVRVKLHQP-EVQSGSYAFSYVYDRAVDMDIDYEGGILKVEDFE 356
Db 332 SSSEVCKAEAAKCYGAVPAPSPLSNIEMGYSEYWYTHDV-----LGQGYDAENIA 387
QY 357 RKAREVCD-----NLENFTSGSP-----FLCNDLSYITALLKDGFGFADST--VLQ 400
Db 388 KTKQYCSKRWSTIQAESKKQLYPRADBEERLRTQCFKSAWITSVLHDGFS-VDKTNKFKQ 446
QY 401 LTKKVNNIETGHWALGA-TFHL 420
Db 447 SVSTIAGQEQVQWALGAMIYHM 467

RESULT 11
ENP1_MOUSE STANDARD; PRT; 510 AA.
ID ENP1_MOUSE
AC P55772;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 1 (EC 3.6.1.5)
DE (NTPDase) (ecto-ATP diphosphohydrolase) (ATPDase) (lymphoid cell
DE activation antigen) (ecto-apyrase) (CD39 antigen).
GN ENTPD1 OR CD39.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95015846; PubMed=7930580;
RA Maliszewski C.R., Delespesse G.J.T., Schoenborn M.A., Armitage R.J.,

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/>) or send an email to license@sib-sib.ch.

EMBL; AF005940; AAB62382.1; -
InterPro; IPR000407; GDAI_CD39_NTPase.
Pfam; PF01150; GDAI_CD39; 1.
PROSITE; PS01238; GDAI_CD39_NTPASE; 1.
KW Hydrolase; Transmembrane; Antigen; Glycoprotein; Calcium; Magnesium.
CYTOPLASMIC (POTENTIAL).
DOMAIN 1 16
FT TRANSMEM 17 37
FT POTENTIAL.
FT EXTRACELLULAR (POTENTIAL).
FT DOMAIN 38 481
FT TRANSMEM 482 502
FT POTENTIAL.
FT DOMAIN 503 513
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 73 73
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 227 227
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 245 245
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 307 307
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 336 336
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 373 373
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 460 460
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 97 97
FT CONFLICT 101 103
FT CONFLICT 101 103
FT CONFLICT 464 464
FT CONFLICT K -> N (IN REF. 2).
FT CONFLICT INV -> CGF (IN REF. 2).
FT CONFLICT K -> V (IN REF. 2).
SQ SEQUENCE 513 AA; 58113 MW; 20FE98F27B6D2F96 CRC64;

Query Match 16.7%; Score 375; DB 1; Length 513;
Best Local Similarity 26.7%; Pred. No. 2.9e-22; Indels 108; Gaps 19;
Matches 130; Conservative 72; Mismatches 177;

QY 8 VFVFLVWVSCVSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMPAGSGTGRHIVYTF 67
DB 17 ILSILGFSCILAVIA-----LLALGLTQNKALPENVK----FGVLDAGSSHTSLYIVRW 67
QY 68 VQKMPGQPLLEGEVDFDVVK-PGLSAFVDQPKQAGTVOGLLEVAKDSIPSEHWKTPV 126
DB 68 PAKKENDTGVTQIEESNVKPGISGFAKKNVINEVLYTACWERAQKVPISQIMSTPVY 127
QY 127 LKATAGLRLL-PEHK--AKALLFEVKEIFRKSPFLVPKGSVSMITQDEGIPAWVTNVL 183
DB 128 LGATAGWLLRMENQMAKILAAVASSISEVFP--DFQGARIISGQEGAYGWTNVL 185
QY 184 TQOL-----HGRQSTVTGLDGAQSTQITLPOPEKTLQTPRGYLTSPMP 231
DB 186 LKFTQKLSWFLNPKPKDQTYGALDLGGASTQITFPVQNETT--ESPNNL-YPKLY 242
QY 232 NSTYKLYTHSYLGFGLKAARLATLGALETGTG----- 265
DB 243 GKNSYVTHSYFLCYGKQQLLOKL-ALGLQTINGLIHPCPSRMYRKIKMSVLNBGSPCT 301
QY 266 --HTPSACLPRNLEAWIPFGVKYQYQNGEVEGVPFCVAELRV-----V 311
DB 302 KRHEINLSFVP-----LVDIETIRGAGN-----FORQSIQLFTSYCPYSSCSF 347
QY 312 RGKLHOPESVQSGSEVAVSYVYDRAVDTMDIDYKGGILKVEDPERKAREVC----- 363
DB 348 NGVFLPLHGGQGFASFTYV-----MEFLMTSEESVSVEQLTEKLEECQAQRWEVQ 401
QY 364 ----DNLENFTSGSPFLCMDLISYITALLKQGF-ADS-TVLQLTQKVNNTIETGALCAT 417
DB 402 KNFGEVKEKLYSE---YCPSTGYILVLLINGHYHFTAESGWNHFNPNKVRSTDVGTGLGM 458
QY 418 PHLLQSL 424
DB 459 LNLTKNI 465

RESULT 14
YNDI_YEAST
ID_YNDI_YEAST STANDARD; PRT; 630 AA.
AC P40009;
DT 01-FEB-1995 (Rel. 31, Created)

103 TVQGLLEVAKDSIPRSHKKTPVVLKATAGLRLLPEHKAKALLPEVKEIPRKGFPL-VPK 161
DB 84 YLTPLRFAEEHIPVEQGEGLLIPATGWRLLPEAQKDAIKVQLNGLKSVTALRVSD 143
QY 162 GSVSMTQDGEIFAWVTNFLTQGLGHROETVGTLDLGGASTQITLPOPEK----- 215
DB 144 SNRIIDGAWEGISWIAVNTILGRFENDSKVGMIDMGASVQIAFEIANEKESYNGG 203
QY 216 TLEQTPRGVLTSEFENSTYKLYTHSYLGFGLKAARLATLGALETGTGHTPSACLPR 275
DB 204 NYVSEINLSIETNEYK--YKLYSTFFLFGYGANGLKKYENSLVKSGNS-----NDSCSPR 257
QY 276 MLEAEWIFGVKRYQYQGNQGEVGEPEPCYAEVLRVVRGKLHOPE----- 319
DB 258 GLNR--LIG-----EFTWANGTGE--NDVCLAQVSSLI-GDKAQPSCPKNPTCFLRNVIAPSV 308
QY 320 EVQRGSFYAFS--YYVDRAVDTMDIDYKGGILKVEDPERKAREVC-----DNLENFTSGSP 374
DB 309 NLSTVOLYGFSEYWTTS-----NFGSGGEVHYQKFTDEVKTKCKQWMDNIQGFGRNE 362
QY 375 F-----LCMDLSYITALLKQGFADST--VLQTLTKKVNIEFTGNALGA----- 416
DB 363 FPNADIERLGTNCFKAAWTSVLHDGFN-VDKTKILFQSVLKIAGEEMQALGAWLYHSK 421
QY 417 --TFHLLQSLGIS 427
DB 422 DLKFNLEQLLEVA 434

RESULT 13
ENPI_BOVIN STANDARD; PRT; 513 AA.
ID ENPI_BOVIN
AC O18956;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 1 (EC 3.6.1.5)
DE (NTPase) (Ecto-ATP diphosphohydrolase) (ATPDase) (Lymphoid cell
DE activation antigen) (CD39 antigen) (Ecto-apyrase).
GN ENTPD1 OR CD39.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Aortic endothelium;
RA Chang A.S., Garcia R.L., Chang S.M., Schilling W.P.;
RA Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RL [2]
RP SEQUENCE OF 97-103; 123-133; 136-140; 145-168 AND 459-471.
RN TISSUE=Aorta;
RX MEDLINE=97115858; PubMed=8955160;
RA Kaczmarek E., Kozlak X., Sevigny J., Siegel J.B., Anrather J.,
RA Beaudoin A.R., Bach F.H., Robson S.C.;
RA "Identification and characterization of CD39/vascular ATP
RA diphosphohydrolase";
RL J. Biol. Chem. 271:33116-33122(1996).
CC -!- FUNCTION: In the nervous system, could hydrolyze ATP and other
CC nucleotides to regulate purinergic neurotransmission. Could also
CC be implicated in the prevention of platelet aggregation.
CC Hydrolyzes ATP and ADP equally well.
CC -!- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 phosphate.
CC -!- COFACTOR: Requires calcium and magnesium.
CC -!- SUBUNIT: Homodimer; disulfide-linked (by similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: Belongs to the GDAI / CD39 NTPase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use.

DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Golgi apyrase (EC 3.6.1.5) (ATP-diphosphatase) (Adenosine
DE diphosphatase) (ADPase) (ATP-diphosphohydrolase) (Golgi nucleoside
DE diphosphatase)
GN YND1 OR YER005M.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RN SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN=S288C;
RC STRAIN=93340091; PubMed=10409709;
RX Gao X.D., Kaigrodov V., Jigami Y.,
RT "YND1, a homologue of GDA1, encodes membrane-bound apyrase required
RT for Golgi N- and O-glycosylation in Saccharomyces cerevisiae";
RL J. Biol. Chem. 274:21450-21456 (1999).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RC MEDLINE=97313264; PubMed=9169868;
RA Dietrich F.S., Mulligan J.R., Hennessy K.M., Yelton M.A., Allen E.,
RA Araujo R., Aviles E., Berno A., Brenan T., Carpenter J., Chen E.,
RA Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,
RA Hunick-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,
RA Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P.,
RA Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,
RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;
RA "The nucleotide sequence of Saccharomyces cerevisiae chromosome V";
RL Nature 387:78-81 (1997).
CC -|- FUNCTION: Catalyzes the hydrolysis of phosphoanhydride bonds of
CC nucleoside tri- and di-phosphates. Has equal high activity toward
CC ADP/ATP, GDP/GTP, and UDP/UTP and approximately 50% less toward
CC CDP/CTP and thiamine pyrophosphate. Has no activity toward GMP.
CC Required for Golgi glycosylation and cell wall integrity.
CC -|- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 phosphate.
CC -|- PATHWAY: Glycosylation.
CC -|- SUBCELLULAR LOCATION: Golgi; membrane-bound.
CC -|- SIMILARITY: Belongs to the GDA1 / CD39 NTPase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AF203695; AAP17573.1; -;
DR EMBL: U18778; AAB64538.1; -;
DR PIR: S50463; S50463.
DR GenOnline: 139085; -;
DR SGD: S0000807; YND1.
DR GO: GO:0017110; P:nucleoside diphosphatase activity; IDA.
DR GO: GO:0006486; P:protein amino acid glycosylation; IMP.
DR InterPro: IPR000407; GDA1_CD39_NTPase.
DR Pfam: PF01150; GDA1_CD39; 1.
DR PROSITE: PS01238; GDA1_CD39_NTPASE; 1.
KW Hydrolase; Transmembrane; Golgi stack.
FT DOMAIN 1 500 LUMENAL (POTENTIAL).
FT TRANSMEM 501 517 POTENTIAL.
FT DOMAIN 518 630 CYTOPLASMIC (POTENTIAL).
FT SEQUENCE 630 AA; 71851 MW; 02P8D24A78212544 CRC64;
Query Match 16.5%; Score 371; DB 1; Length 630;
Best Local Similarity 26.0%; Pred. No. 8.1e-22;
Matches 116; Conservative 75; Mismatches 164; Indels 92; Gaps 17;
49 YGIMFDAGSTGRTHVYTF-----VQMPQQLPILSGE---VFD5VKPGLSA 92
10 FGIVIDAGSGSRHVFVKQDTESLLHATNODSQSILQSVPHIQEKDWTF-KLNPGLSS 68

QY 93 FVDPQKQAGT-VQGLLEKADSIIPRSHWKKTPVVLKATAGLRLLPEHAKALLPEV-KE 150
DB 69 FKKQDQAYKSHIKELDFAKNIIIPESHSSCPVFIQATAGNRLLPQDIQSSILDLGCG 128
QY 151 IFRKSPFLVP--KGSVSIQMDQDEGIFAWTVNFTLQGLHGRQE-----TVGTLDLGA 203
DB 129 LKHPAEFLVEDCSAQIQVIDGETGLYGLWGLNLYLYGHFNNDYNPEVSDHFTFGMDWGA 188
QY 204 STQIIFLPQFETLE-----QTPRGVLTSPFENSTYKLYTHSYLGFGLKAARL 252
DB 189 STQIAFAPHDSGEIAHRDDIATIFLRSVNGDLQKWDVFSV-----WLGFGANQARR 241
QY 253 ATLGAL-----ETEGTDGHTFRSACLPRLWEAEWIGGVKVGKQGEVGFPE 301
DB 242 RYLAQLINTLPENTDYNDEDFSTRNLNDPCMPRGSSDTDFEKDTIFHAGSNGYQCTK 301
QY 302 PCVAEVLV-----VVRGKLHQP-EVQSGSFYAFSVYVDRAVTDMDVKEKGILKV 352
DB 302 STYPLLLKNMPCDDEPCLFNGVHAPRIDPANKFICTSEYTWYANDV-----FKLGSEYNF 357
QY 353 EDPFKAREVCQN-----LENFTSG-----SPFL---CMDSLYITALLKDGFGPA--- 394
DB 358 DKFSKSLREFCNSNWTQILANSKGVVNSIPENFLKDACFKGNWVLNHLHEGDFMDRIDV 417
QY 395 -----DSTVHLQTKKNNIETGALG 415
DB 418 DAENVNDRFLFQSVKEKVERELSWTLG 444
RESULT 15
ENPL_PIG STANDARD; PRT; 510 AA.
AC QSMYU4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 1 (EC 3.6.1.5)
DE (NTPase) (Ecto-ATP diphosphohydrolase) (ATPase) (Lymphoid cell
DE activation antigen) (Ecto-apyrase) (CD39 antigen).
GN ENTPI1 OR CD39.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=Aortic endothelium;
RC MEDLINE=20325090; PubMed=10866813;
RA Lemmens R., Vanduffel L., Kittel A., Beaudoin A.R., Benrezzak O.,
RA Seigny J.;
RA "Distribution, cloning, and characterization of porcine nucleoside
RA triphosphate diphosphohydrolase-1";
RL Eur. J. Biochem. 267:4106-4114 (2000).
RN [2]
RN SEQUENCE OF 202-220.
RC TISSUE=Pancreas;
RC MEDLINE=97115858; PubMed=8955160;
RA Kaczmarek E., Kozlak X., Seigny J., Siegel J.B., Anrather J.,
RA Beaudoin A.R., Bach P.H., Robson S.C.;
RA "Identification and characterization of CD39/vascular ATP
RA diphosphohydrolase";
RL J. Biol. Chem. 271:33116-33122 (1996).
CC -|- FUNCTION: In the nervous system, could hydrolyze ATP and other
CC nucleotides to regulate purinergic neurotransmission. Could also
CC be implicated in the prevention of platelet aggregation.
CC Hydrolyzes ATP and ADP equally well.
CC -|- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 phosphate.
CC -|- COFACTOR: Requires calcium and magnesium (by similarity).
CC -|- SUBUNIT: Homodimer; disulfide-linked (by similarity).
CC -|- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -|- TISSUE SPECIFICITY: Highest expression found in vascular
CC endothelium, smooth muscle, spleen and lung.
CC -|- PTM: CLEAVED IN TWO POLYPEPTIDES THAT SEEM TO STAY TOGETHER BY

